

# Phonetica

Internationale Zeitschrift für Phonetik

International Journal of Phonetics

Journal international de phonétique

*Editor:*

E. Zwirner *Münster in Westfalen*

*Coeditores:*

R. Avanesov	<i>Moskva</i>
E. Blanquaert	<i>Gent</i>
S. K. Chatterji	<i>Calcutta</i>
W. Doroszewski	<i>Warszawa</i>
E. Fischer-Jørgensen	<i>København</i>
I. Fónagy	<i>Budapest</i>
D. B. Fry	<i>London</i>
P. L. Garvin	<i>Los Angeles, Calif.</i>
H. Glinz	<i>Kettwig</i>
R. Jakobson	<i>Cambridge, Mass.</i>
K. Jimbo	<i>Tokyo</i>
M. Joos	<i>Madison, Wisc.</i>
B. Malmberg	<i>Lund</i>
A. Martinet	<i>Paris</i>
Sh. Morag	<i>Jerusalem</i>
R. Olesch	<i>Köln</i>
G. E. Peterson	<i>Ann Arbor, Mich.</i>
K. L. Pike	<i>Santa Ana, Calif.</i>
A. A. Reformatkij	<i>Moskva</i>
L. E. Schmitt	<i>Marburg</i>
E. W. Selmer	<i>Oslo</i>
A. Sovijärvi	<i>Helsinki</i>
G. Straka	<i>Strasbourg</i>
F. Trojan	<i>Wien</i>
L. Zabrocki	<i>Poznań</i>

*Secretarius:*

W. Bethge *Münster in Westfalen*



S. KARGER BASEL - NEW YORK

# INHALT – CONTENTS – SOMMAIRE

Vol. 7, No. 2/3 (1961)

*Originalarbeiten – Original Papers – Travaux originaux*

G. E. PETERSON and N. P. MCKINNEY (Ann Arbor, Mich.):	
The Measurement of Speech Power – <i>Die sprachlich-akustische Kraftmessung</i> –	
La mesure de l'énergie de la parole . . . . .	65
B. MALMBERG (Lund):	
Phonèmes labio-vélaires en espagnol? – <i>Labio-velare Phoneme im Spanischen?</i> –	
Labio-velar Phonemes in Spanish? . . . . .	85
W. KALLENBACH und H.-J. SCHROEDER (Braunschweig):	
Zur Technik der Tonbandaufnahme bei Sprachuntersuchungen – <i>Tape</i>	
<i>Recording Techniques in Voice Research</i> – Remarques sur la technique d'enregist-	
rement d'études phonétiques . . . . .	95
A. ROSETTI (Bucarest):	
Sur le problème de la syllabe. Réponse à M. B. Hála – <i>Zum Problem der Silbe.</i>	
<i>Antwort an B. Hála</i> – On the Problem of the Syllable. Answer to B. Hála	109
I. LEHISTE (Ann Arbor, Mich.):	
Some Acoustic Correlates of Accent in Serbo-Croatian – <i>Akustische Korrelate</i>	
<i>des Akzents im Serbokroatischen</i> – Quelques corrélats acoustiques de l'accent	
serbocroate . . . . .	114
Y. OCHIAI (Nagoya):	
Mémoire sur l'analyse de la qualité des voyelles – <i>Zur Analyse der Vokal-</i>	
<i>qualitäten</i> – Note on Analysis of Vowel Quality . . . . .	148
LIBRI . . . . .	160
VARIA . . . . .	190

Einem Teil der Auflage dieses Heftes liegt ein Prospekt der Firma ELCO AG,  
Basel, bei.

## General Information

**Phonetica** are published in issues of 64 pages. 1 volume consists of 4 number and costs sFr. 56.– (postage incl.). 2 volumes appear annually. All manuscripts should be addressed to the Institut für Phonometrie, Steinfurter Strasse 107, Münster i. Westfalen (Germany). Books for review as well as enquiries concerning subscriptions and advertisements should be sent to the publishers.

Contributors will receive 50 copies of their papers free of charge in lieu of remuneration. Extra copies, if desired, should be ordered with the return of the first proofs and will be supplied at a special rate.

Articles will be printed in either German, English or French, with summaries in all three languages.

Rules for the preparation of bibliographies (with examples) as well as lists of abbreviations of the most frequently quoted Journals may be obtained from the publishers.

## Publikationsbedingungen

**Phonetica** erscheinen in Heften zu 64 Seiten. 4 Hefte bilden 1 Band zum Preis von sFr. 56.– (inkl. Porto). Jährlich erscheinen 2 Bände. Alle Manuskripte sind zu senden an das Institut für Phonometrie, Steinfurter Straße 107, Münster i. Westfalen (Deutschland). Rezensionsexemplare sowie Zuschriften, Abonnements und Inserate betreffend, sind an den Verlag zu richten.

Die Mitarbeiter erhalten anstelle eines Honorars 50 Sonderdrucke kostenfrei. Weitere Sonderdrucke, die bei der Rücksendung der ersten Korrektur zu bestellen sind, liefert der Verlag gegen Berechnung.

Publikationssprachen sind Deutsch, Englisch und Französisch. Jede Originalarbeit erhält eine Zusammenfassung in diesen drei Sprachen.

Anleitungen zur Abfassung von Bibliographien (mit Beispielen) sowie Abkürzungslisten der wichtigsten Zeitschriften können vom Verlag angefordert werden.



Communication Sciences Laboratory, Ann Arbor, Michigan

## **The Measurement of Speech Power\***

By GORDON E. PETERSON and NORRIS P. MCKINNEY

Speech power measurement has been an essential procedure in the development of electrical communication systems. However, the measurement of speech power and other closely related functions of the acoustical speech wave for the purpose of linguistic analysis has until recently received relatively little attention. Speech power, loudness, and phonetic stress are often considered to be closely related or even synonymous. Actually, however, an understanding of the relation between speech power as a function of time and linguistic judgments of stress is an essential problem in experimental linguistics. By "linguistic judgments" we refer to the responses of observers when they are reacting to speech as a form of communication rather than as a sequence of abstract sounds.

There is an extensive literature on the relationship of the loudness of various types of elementary acoustical signals to the measurement of their powers. These signals include single and multiple pure tones, frequency bands of white Gaussian noise, etc. It is not at all clear, however, that the results of such studies on simple acoustical signals can be extended to predict the "loudness" with which speech is perceived. Although little is known about the relationships among speech power, loudness, and stress, it is clear that stress perception is to some extent based upon speech power<sup>9</sup>. Stress judgments are also affected by the fundamental frequency of voiced sounds, duration, vowel quality, and voice quality<sup>5, 9</sup>.

Obviously, then, no simple power measurement on the acoustical waveform will specify linguistic judgments concerning speech. Because the relationship of speech power to stress is complicated,

---

\* This research was supported by the Office of Scientific Research of the United States Air Force under contract AF 49(638)-492.

and because there are so many other factors which also affect the perception of stress, it is plausible that some measure of the power involved in the production of speech may be more closely related to judgments of linguistic stress than are measurements of acoustical speech power. Speech production power may be considered the product of the pressure and volume velocity of the air entering or being emitted from the vocal mechanism during speech. The extent to which measures of speech production power will correlate with listener judgments of linguistic stress, however, is as yet unknown.

In many applications only the acoustical signal is available and physiological measures are impractical. Automatic speech recognition, for example, is an instance in which stress and other linguistic interpretations must be based primarily upon acoustical data. It does not seem unreasonable to attempt to make such interpretations, since obviously they can be made by a listener without reference to physiological observations.

The above considerations indicate the need in speech research for circuitry suitable for the measurement of speech power. In general, an instrument for the measurement of acoustical speech power or of some closely related function should be sensitive to changes in the amplitude of a given input waveform, and should be relatively insensitive to changes in the frequency distribution of the power. It should present relevant information and should exclude extraneous components from the output. The remainder of this paper is concerned with these and other design considerations of a circuit for the measurement of speech power for linguistic studies. *Fant* has previously described a similar circuit which he calls an "intensity meter" (3, pp. 334-339).

Let  $p(t)$  symbolize the electrical representation (current or voltage) of the acoustic pressure wave of the speech which is the input to a speech power measuring circuit. The first consideration is whether a weighting of the spectrum should be performed on the input wave. Such a weighting would be performed by a circuit represented by the block at the left in figure 1. The output of the circuit may be represented by  $p'(t)$ .

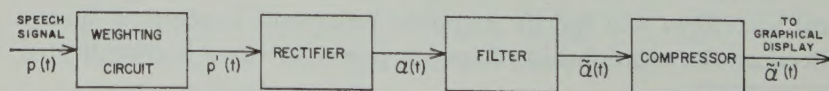


Fig. 1. Block diagram of a circuit for the measurement of speech power.



If such a weighting is to be performed, it would seem most reasonable to base it on the well known equal-loudness contours of audition. Since these contours vary with both frequency and intensity level, the input-output characteristic of the corresponding weighting circuit must vary as a function of both the frequency and the level of the individual spectral components. A device which could perform such an operation would, of course, be exceedingly complicated. There are many reasons, however, to question whether the use of any such weighting is indicated at the present time, or would be of any particular assistance in deriving measures related to judgments of linguistic stress. As indicated above, it is not known that the results of tests with elementary acoustic signals are applicable to the interpretation of speech. Also, at the levels normally of interest in speech and throughout the frequency range involving maximum contribution to speech intelligibility, the equal-loudness contours are relatively flat and have approximately equal spacings. In this paper and in the circuitry to be described, such weighting is considered impractical and possibly irrelevant; hence  $p'(t)$  is equal to the input signal,  $p(t)$ .

If a constant pattern of acoustical radiation is assumed, power and intensity measures have a direct correspondence. The term *power* will be employed throughout the present paper as a measure of energy per unit time. Electrical power measurements are obtained from the square of the voltage or current. If  $p'(t)$  is the instantaneous voltage, then  $[p'(t)]^2$  is proportional to the instantaneous power. This operation is normally accomplished by taking the absolute value of the linear waveform, and by squaring the result. It is performed by the rectifier circuit, represented by the second block in figure 1. The output is a nonlinear function indicated by  $a(t)$ .

Most rectifier elements have an output  $a(t)$  approximately proportional to the absolute-value function  $|p'(t)|$  and most of the remaining ones have an output approximately proportional to the square function  $[p'(t)]^2$ . There seems to be little reason to consider other rectifier characteristics for application to speech power measurement. The major design problems associated with the rectifier, then, are to determine for both the square and absolute-value functions the extent to which each of these mathematical ideals can be achieved, the relative difficulty of achieving an approximation to the ideal, and the relative value of doing so. On the one hand, since the definition of power requires the square of the acoustic pressure

function, the use of a poor approximation to the ideal square function would reduce the accuracy of the system for power measurement. On the other hand, a "linear rectifier" (i.e. one which approximates the absolute-value function) may easily be made to approximate the ideal over a much wider range of input voltage levels than can a "square-law rectifier". Thus the difference observed between a linear rectifier output and a square-law rectifier output when a representative sample of speech is applied at the input merits consideration.

While the response of the human to power fluctuations in such complex signals as speech has not been well defined, it is clear that a speaker's control of amplitude variations or a listener's interpretation of them is very much slower than the rates of variation present in the instantaneous output of the rectifier. Expressed differently, the information rate associated with the rectifier output, which is almost as great as that associated with  $p(t)$ , is many times the information rate which the human is capable of controlling or perceiving. From the experimental view it is therefore appropriate that the psychophysicist or the experimental linguist seek a limited number of independent amplitude measures within an utterance which he may relate to listener judgments about those utterances. Thus a slowly varying function  $\bar{a}(t)$  is desired, and from this function values which appear most relevant to the interpretation of the speech signal may be selected. The third block of figure 1 represents the filter circuit which is required for smoothing the rapidly fluctuating output from the rectifier. The output of the filter circuit,  $\bar{a}(t)$  is an averaged or filtered function which the circuit derives from its input  $a(t)$ .

The filter, then, averages the rapid fluctuations of the rectifier output over a short time interval and presents as its output the slowly varying  $\bar{a}(t)$  for interpretation. The problem associated with this section of the speech power circuit is, of course, the specification of the filter function  $f(\tau)$ . This includes the specification of the manner in which the filter is to respond and the specification of the rate of the filter response to rapid changes at the input.

Finally, since the range of values of  $\bar{a}(t)$  which may be significant to the psychophysicist or experimental linguist is very large, it may be desirable to compress the representation of some of the larger input signal levels into a smaller portion of the overall range of the output than occurs when  $\bar{a}(t)$  is presented linearly. Such a compression makes it possible to observe the lower intensity level vari-



ations more accurately. These output level adjustments are accomplished by a compressor represented by the right-hand unit shown in figure 1. The output of this unit might be a logarithmic function of its input, for example, in which case the output could be calibrated linearly in decibel units.

The following sections of the paper will review in sequence some considerations involved in the design of a rectifier, a filter, and a compressor for the measurement of speech power.

### *Rectification*

The optimum rectifier characteristic for our purposes is that which causes the smoothed function  $\bar{a}(t)$  to correspond most nearly to the perceived stress of the speech sound represented by the rectifier output. Present knowledge of the relationships between perceptual categories of speech and measures of the signal  $p(t)$  is insufficient to dictate the most appropriate rectifier characteristic. It is interesting to note that *Fletcher*<sup>4</sup>, with only a brief word of justification, devotes a chapter to a discussion of acoustical speech powers. Altogether he defines six different measures of power for describing speech sounds (<sup>4</sup>, pp. 68–69).

The two candidate rectifier characteristics, the square and the absolute-value, both have the advantage of being elementary functions. The square function is at every instant proportional to the instantaneous power; its average is proportional to the average power. Average power is the summation of the powers of the spectral components whose periods are short compared to the averaging time; average power does not change when only the phase relationships of the components of the wave are changed. Average power measures are thus relatively insensitive to phase shifts of the spectral components with respect to each other, and this insensitivity to phase is true also, to a first approximation, of the auditory perception of sound (<sup>10</sup>, pp. 1024–1026).

The theoretical advantages of the square function, however, are accompanied by considerable practical difficulty in designing circuitry to operate according to a square-law characteristic. No single circuit element is to be found which combines the needed speed of response and dynamic range. Very good approximations to a square-law response may be obtained over a limited range by means of a diode function generator, such as is found in the Ballantine Model

320 True Root-Mean-Square Voltmeter. We have employed this instrument as a square-law circuit in an intensity meter such as is depicted in figure 1. Since the range of measurement is restricted, however, the power function of continuous speech with a diversity of levels cannot be recorded in a single data analysis.

As indicated previously, "linear rectification", which forms a signal proportional to the absolute value of the input speech wave, is much easier to achieve instrumentally than is square-law rectification. Rectifier circuits employing ordinary semiconductor diodes can be designed to approximate closely the absolute-value function over a 75 decibel dynamic range of the input. This is a considerably greater range than can be obtained with square-law rectification. A second consideration is that the output amplitude of the linear rectifier covers the same decibel range as the input, while the output amplitude of the square-law rectifier covers twice the decibel range of the input. This consideration is relevant because the range of nonlinear circuit elements eligible for use in the compression circuit, whose input is the rectifier output, is rather limited.

The average of the output of a linear rectifier, however, varies with changes in the phase of spectral components with respect to each other, even when the amplitudes of the components remain constant. This is generally in contrast with both human perception of the loudness of elementary acoustic signals, and with the output of the ideal square-law rectifier.

In order to examine quantitatively the error introduced by using a linear rectifier rather than a square-law rectifier, an experimental comparison was made between the output of a circuit employing linear rectification and the output of a similar circuit employing square-law rectification, with identical inputs to both. The differences occurring between the outputs of the two systems were measured and tabulated.

A block diagram of the equipment used for the experimental comparison is shown in figure 2. Since the analyses were also needed for another experiment<sup>8</sup>, the speech data input to the two systems were Serbian words in sentence frames uttered by a male speaker DK. They were tape-recorded on an Ampex recorder in a sound treated room. The utterances were then reproduced to both inputs of the two rectification systems. In the linear channel, the signal was amplified and fed to a full-wave semiconductor diode rectifier system. The rectified waveform was first smoothed by a 76 cps



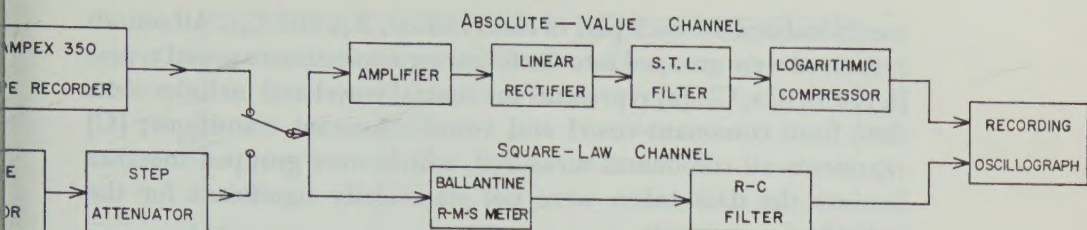


Fig. 2. Equipment used for the experimental comparison of square-law and absolute-value rectification of speech signals.

low-pass filter (see next section) and then compressed to a logarithmic scale (as described later in the section on compression). In the square-law channel, the signal from the tape recorder was applied to the input of the True Root-Mean-Square Voltmeter, the filter capacitor of which had been removed from the circuit. The output current of the meter, which is proportional to the square of the instantaneous input voltage, was smoothed by a low-pass, single time-constant, resistance-capacitance filter. The outputs of the filters of the two channels were applied to the inputs of the two channels of a Sanborn two-channel hot-wire recording oscillograph. Although the filters in the two channels were not identical, they were both low-pass filters having their cutoff frequencies high enough so that the recording oscillograph response dominated the low-frequency response. A sine-wave signal input was varied in one decibel steps to calibrate both channels and the oscillograph simultaneously. The R-M-S channel was limited to a ten decibel range at a single time, but the range was changed for different parts of the recording.

At various points along the time axis, the value of the ordinate was read from each of the two recording oscillograph traces. Usually on such sounds as vowels the smoothed power was read where a maximum occurred. In other cases the value was read at a minimum or at some other identifiable point. These raw data were converted to levels in decibel units by means of a curve based on the oscillograms obtained with the calibrating signal as input to the systems.

Let  $R_{jk}$  be the level indicated by the square-law channel and  $L_{jk}$  be the level indicated by the linear rectifier channel for the  $k^{\text{th}}$  token of sound type  $j$ . The difference (in decibels)

$$D_{jk} = R_{jk} - L_{jk}$$

between the levels indicated by the outputs of the two data channels

was calculated for each pair of observations,  $R_{jk}$  and  $L_{jk}$ . All sounds measured were grouped into the following approximate sound types: [i, e, a, o, u, ə, C]; [ə] represents the central vowel and includes some data from consonant-vowel and vowel-consonant transitions; [C] represents all consonants measured, which were grouped together because the data taken were not statistically significant for the individual consonants.

The mean difference

$$S_j = \frac{\sum_k D_{jk}}{n_j}$$

was calculated for each sound type, where  $n_j$  is the number of tokens measured of sound type  $j$ . The 95% confidence interval was also calculated for each mean based on the assumption of a normal distribution of the difference  $D_{jk}$  for each sound type about the sample mean  $S_j$ . A chi-squared test of this assumption yields  $\chi^2 = 10.9$  with 12 degrees of freedom. The sample standard deviation of  $D_{jk}$  was also calculated for each sound type  $j$  and is shown in figure 3, together with the sample means and their confidence intervals.

The principal conclusions which may be drawn from the limited data just reviewed are that if the square-law rectifier is taken as the ideal, then 1. the error in the speech power measurement made by means of a linear rectifier will seldom exceed two decibels, and 2. the error will average in the order of one decibel and will always be positive. In view of the limitations of the human to distinguish sound level differences, even under controlled conditions, it appears that linear rectification or some closely related function will provide a satisfactory measurement of speech power for many purposes of linguistic analysis. Where desired, a square-law device appears practical for the measurement of ranges narrower than the total involved in speech.

### *Filtering*

A graphical representation of the original speech signal  $p(t)$  or of an unsmoothed function of that signal, e.g.  $a(t)$ , is usually of little use to the experimental linguist. Although the information he desires is contained in such a representation, it is in such a complicated form that it is essentially unavailable to him. The function of any speech parameter extraction device (e.g. a speech power



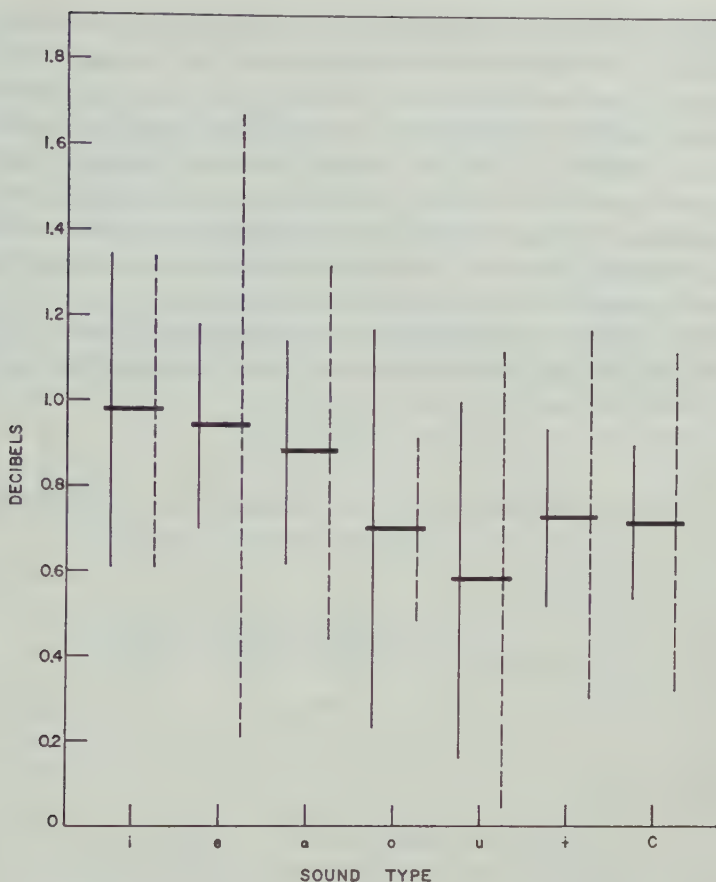


Fig. 3. The sample mean (horizontal bar), standard deviation (dashed line), and 95% confidence interval for the mean (solid line) for the difference between square-law and absolute-value measurements on a speech sample.

circuit) is to form from the speech wave a signal which has a highly reduced information rate, but which retains the desired speech parameter measure (e.g. speech power). The principal information reducing component of such a parameter extraction device is ordinarily a filter which has a bandwidth much less than that of the original speech wave. A major filter design problem is the selection of the filter bandwidth. As the bandwidth is increased the linguist is increasingly assured of having all of the relevant information available to him, but he also has an increasing number of decisions to make in interpreting the data.

A second filter design problem is the compromise that must be made between the time domain response and the frequency domain response. The frequency-domain and time-domain considerations are closely related to the desire to make the output trace of the system correspond as closely as possible to auditory judgments of the acoustic signal<sup>1</sup>. Our present knowledge of the response of the auditory system to acoustic inputs, and to speech inputs in particular is so incomplete, however, that we must depend considerably upon generally accepted criteria of good filter design.

The process of filtering the instantaneous speech power  $a(t)$  to obtain the average speech power  $\bar{a}(t)$  may be described by the convolution integral equation

$$\bar{a}(t) = \int_0^t a(u) f(t-u) du$$

where  $u$  is the time variable of integration,  $f(u)$  is the filter time-

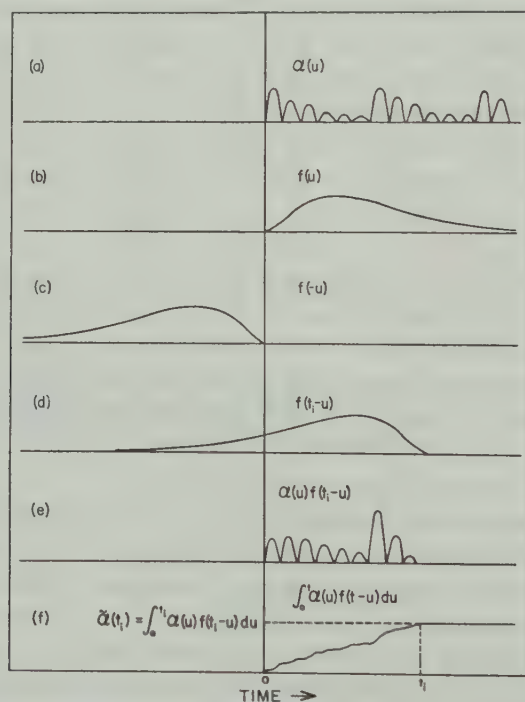


Fig. 4. Filtering as described by the convolution integral. a) waveform to be filtered; b) impulse-response of the filter; c) inverted impulse-response; d) inverted impulse-response with its leading edge advanced to time  $t = t_1$ ; e) product of the waveforms of (a) and (d); f) the point of data obtained by integrating the product waveform of (e) and the integral curve of which the point is a member.



response to a unit impulse input (i.e. impulse-response), and the input  $a(u)$  is zero until time  $t = 0$ . One may consider the action of the filter as the application of a weighting function  $f(-u)$  which moves along the time axis with its leading edge at time  $t$ , as portrayed in figure 4.

A considerable proportion of speech sounds are voiced, and most of the resultant sound pressure waveforms may be assumed to be approximately periodic over intervals of two or three times the fundamental period of the speech wave. This quasi-periodicity makes it desirable for an ordinary filter in this application to have an integrating time of several periods to avoid "ripple" in  $\bar{a}(t)$ . On

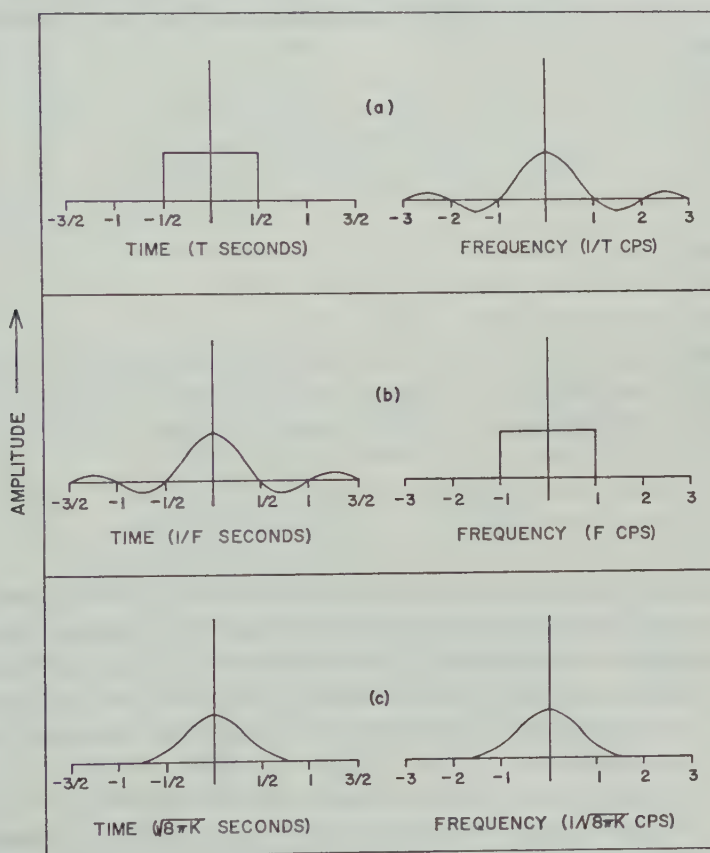


Fig. 5. Time-response (impulse-response) and frequency-response of a) an ideal rectangular time-response filter, b) an ideal rectangular frequency-response filter, and c) an ideal Gaussian filter.

the other hand, some instances of apparent linguistically significant changes in the power level of the speech waveform have been observed within a time interval only three or four voice periods in length. A long integrating time would obscure such changes.

These two considerations taken together make it seem desirable to use the uncommon rectangular time-response filter function  $f(u)$  having an integrating time  $T$  equal to exactly one or two fundamental voice periods. The rectangular time-response of such a filter and the corresponding frequency-response are shown in figure 5a. The integration might be discrete in time, a new output occurring every complete cycle or two of the speech wave. Such a measure could be implemented by an electronic gate. If the integration of the gate output is made over single successive voice periods of duration  $T_n$ , then:

$$\tilde{a}_{T_n}(t) = \frac{1}{T_n} \int_{\sum_{i=0}^{n-1} T_i}^{\sum_{i=0}^n T_i} a(u) du$$

A continuous rectangular filter could be implemented by a delay line. If  $T$  represents a continuous measure of the voice period, the resulting average speech power function would be:

$$\tilde{a}_T(t) = \frac{1}{T} \int_{t-T}^t a(u) du$$

Serious problems arise, however, in the design of electronic instrumentation for measuring this function  $\tilde{a}_T(t)$  because of the large dynamic range, both in the level of  $a(u)$  and in the fundamental frequency of  $p(t)$ . In addition, division or multiplication must be accomplished over the range of the fundamental voice periods encountered.

Furthermore, the instrumentation of the function  $\tilde{a}_T(t)$  requires a signal proportional to  $\frac{1}{T}$ , i.e. proportional to the fundamental frequency. Despite the longstanding interest in a circuit whose output is proportional to voice fundamental, none is as yet available that is sufficiently reliable for such an application. An



added complication is that switching circuits are necessary in such a filter. During the course of this study an estimate was obtained of the order of magnitude of the maximum error in  $\bar{a}(t)$  due to integrating over a time interval longer than the fundamental period. This maximum error was found to be about 3 decibels for an example of  $[a]$  produced with good phonation. The error is considerably increased in the case of more impulsive waveforms, such as that of a laryngealized  $[a^?]$ .

In the frequency domain, one idealization of the filter frequency-response is that it should be constant from zero cycles-per-second up to some frequency just below the lowest voice fundamental frequency encountered. Such a frequency-response is said to be rectangular and is portrayed in figure 5b, together with the associated time-response. This filter would pass all the significant fluctuations in the speech power level and would exclude from its output,  $\bar{a}(t)$ , the "ripple" components which interfere with the interpretation of  $\bar{a}(t)$ . Parameter digitalization for the purpose of automatic speech recognition may place considerable restriction on the allowable ripple in the filter output.

The time response to certain impulsive waveforms, however, such as those associated with plosive consonants and the sharp front of the onset of voicing, must not be badly distorted. Such distortion might easily mask or destroy linguistically significant information. Simple time delay between the input and the output is not troublesome, but the approximate waveshape of the envelope should be preserved. This requires that all components passed by the filter be delayed by approximately the same amount of time. There should be little or no ringing of the filter. It is on the basis of such time domain criteria that approximations to the rectangular frequency response filter are disqualified.

*Klapper and Harris*<sup>7</sup> have shown that a filter having a Gaussian response is the optimum compromise between certain restrictions in the time and frequency domains. Both the impulse-response and the frequency-response of such a filter are Gaussian in shape, as is shown in figure 5c. The ideal Gaussian filter is not physically realizable, but *Klapper and Harris* have shown how it may be approximated to any desired accuracy. They have given data on a filter they constructed having five stages, with two pairs of poles in each stage or ten pole pairs altogether.

A criterion which measures the combined abilities of a filter to

respond quickly to changes in the input and to eliminate ripple components from its output is the product of its bandwidth and its integrating time or rise-time\*. For the Gaussian filter described by *Klapper and Harris* the bandwidth-integrating time product is

$$BW \times t_1 = .343$$

The frequency-response drops off with a large negative slope approaching -60 db per octave at the higher frequencies, but it has a rather gradual "knee" as shown in figure 6. The filter's time response to an impulse function input is shown in figure 7.

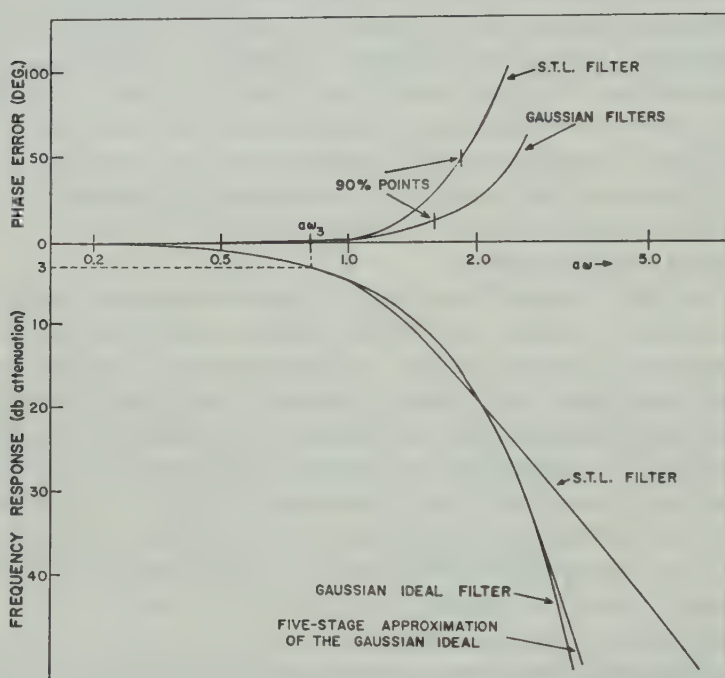


Fig. 6. Frequency-response and phase error-response curves for Gaussian filters and the S.T.L. filter.

*Fant*<sup>2</sup> has described a filter (S.T.L. filter) derived by *Voss*<sup>12</sup> which is used at the Speech Transmission Laboratory in Stockholm

\* The bandwidth,  $BW$ , of a low-pass filter is defined as the frequency at which the amplitude response drops to 3 db below its low-frequency asymptote. The integrating time,  $t_1$ , is defined as the amount of time required for the output response to a step function input to rise from 0.1 of its final value to 0.9 of its final value.



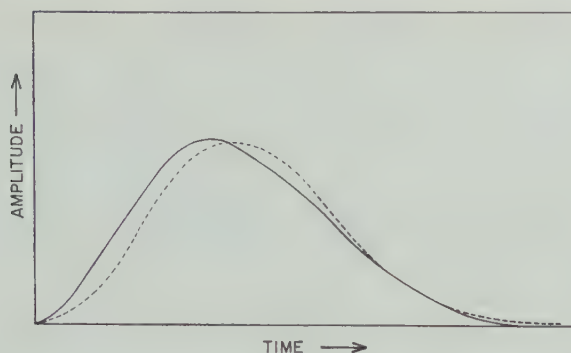


Fig. 7. Filter impulse-responses for the S.T.L. filter (dashed line) and the 5-stage approximation to the Gaussian ideal filter.

and which is considerably simpler than any of the filters mentioned above. It is a symmetrical pi-network that has been phase-compensated to provide a good transient response. The performance of the S.T.L. filter compares well with that of the Gaussian filter. Its complex transfer function has three poles, two of which are a nearly critically damped complex conjugate pair. Its bandwidth-integrating time product is

$$BW \times t_1 = .357,$$

nearly the same as for the Gaussian filter.

Another criterion mentioned by *Klapper and Harris* (although not calculated in their paper) is that of phase-response linearity. Distortion of transients is minimized by having the phase shift as nearly proportional to frequency as possible. The criterion they mention is that the phase-response linearity should extend over more than 90 percent of the area under the frequency-response curve. The phase-response of the Gaussian approximation filter has departed 13 degrees from a linear response at the 90 percent point, while that of the S.T.L. filter has departed 48 degrees (figure 6). This means that transient speech parameter wave-shapes will be somewhat less distorted by the Gaussian filter.

The frequency-response for the S.T.L. filter does not have such a large negative slope at the higher frequencies, but in the neighborhood of the cutoff frequency the two filters are similar. Their frequency-responses do not depart more than  $1\frac{1}{2}$  db from each other until beyond two and one-half times their cutoff frequency, where they are down more than 20 decibels from their low-frequency

responses (figure 6). That the two filters are very similar may also be seen by a comparison of their impulse-responses, given in figure 7. Because of its approximation to a Gaussian ideal filter and because of its simplicity, the S.T.L. filter was chosen for speech power measurements at the Communication Sciences Laboratory.

The cutoff frequency for the filter should be a compromise value which will allow the ripple components to be sufficiently attenuated, and which will still preserve the amplitude and waveform of significant transients. In the speech power circuit constructed, several cutoff frequencies are available, one of which is selected for each data run. If a system is to have only one bandwidth and is to be used in general applications, a good compromise bandwidth is about 50 cps. This is somewhat lower than the lowest fundamental frequency encountered in speech with non-laryngealized phonation, but is high enough to keep the integrating time small and thus allow significant transients to pass.

### *Compression*

Compression of the signal output from the filter is desirable for at least two essential reasons: 1. *Stevens* has found that for abstract sounds  $\log L = 0.6 \log \tilde{a} - K$ , where  $L$  is estimated loudness and  $K$  is a constant<sup>11</sup>, and 2. optimal use is desired of the range of the oscillograph or other display device to which the speech power circuit output is applied. Thus compression according to a logarithmic function appears reasonable. The compressed function may be defined as:

$$\tilde{a}'(t) = K_1 \log \tilde{a}(t) + K_2,$$

where  $K_1$  and  $K_2$  are constants. This function also describes the current-voltage relationship of a semiconductor diode over a considerable portion of its operating range. The circuit of a compressor utilizing the logarithmic current-voltage characteristic of a diode would consist of a diode driven by a source current proportional to the filter output. The diode voltage is applied to an amplifier with a high input impedance, to provide the output of the speech power circuit. Diodes are available with a useable logarithmic range of over 100 db, and the driving and isolation amplifiers needed in connection with their use are entirely feasible. Thus the compressor design is a relatively minor problem.



### *A Note on Interpretation*

According to *House*<sup>6</sup>, there may be considerable variations in speech power output as vowel quality is held constant and the fundamental frequency of the voice is changed. This effect was observed in the use of the above described speech power circuit. In a study to be reported later by *Ladefoged*, a few words were spoken with a long falling pitch. The power level oscillogram was observed in some cases to have one to four rather prominent peaks, although the words were phonated evenly. It was found that these peaks occurred at just the times at which the harmonic frequencies passed through formant center frequencies. This correlation may be seen in figure 8 by comparing the oscillograms shown there with the two narrow band spectrograms for the same sounds. The sounds represented are vowel sounds produced with subjectively constant effort and vowel quality and with falling pitch. The differences in amplitude between the maxima and minima of the intensity curve were not as great as those predicted on the basis of the first-order models described by *House*, but are accounted for in his paper.

The relationship of the fundamental frequency to the frequencies of the formants of highest amplitude thus also influence measures of speech power. In interpreting speech power data for voiced speech sounds, then, the relationship between the glottal spectrum and the supraglottal transmission function for that sound should also be considered.

### *Conclusions*

From the foregoing data and discussion we may conclude the following:

1. The measurement of speech power level is relevant to the study of linguistic stress.
2. The power level may be measured by means of a circuit with a rectifier followed by a filter.
3. It is presently not appropriate to weight the spectrum of the speech signal before it is rectified.
4. Linear rectification is most easily accomplished, and should be satisfactory for most linguistic studies. Square-law rectification, although more difficult, has theoretical advantages.
5. A filter having a response closely approximating the Gaussian ideal adequately meets the relevant filter criteria. A relatively simple

<sup>6</sup> *Phonetica*, Vol. 7, Nr. 2/3, 1961

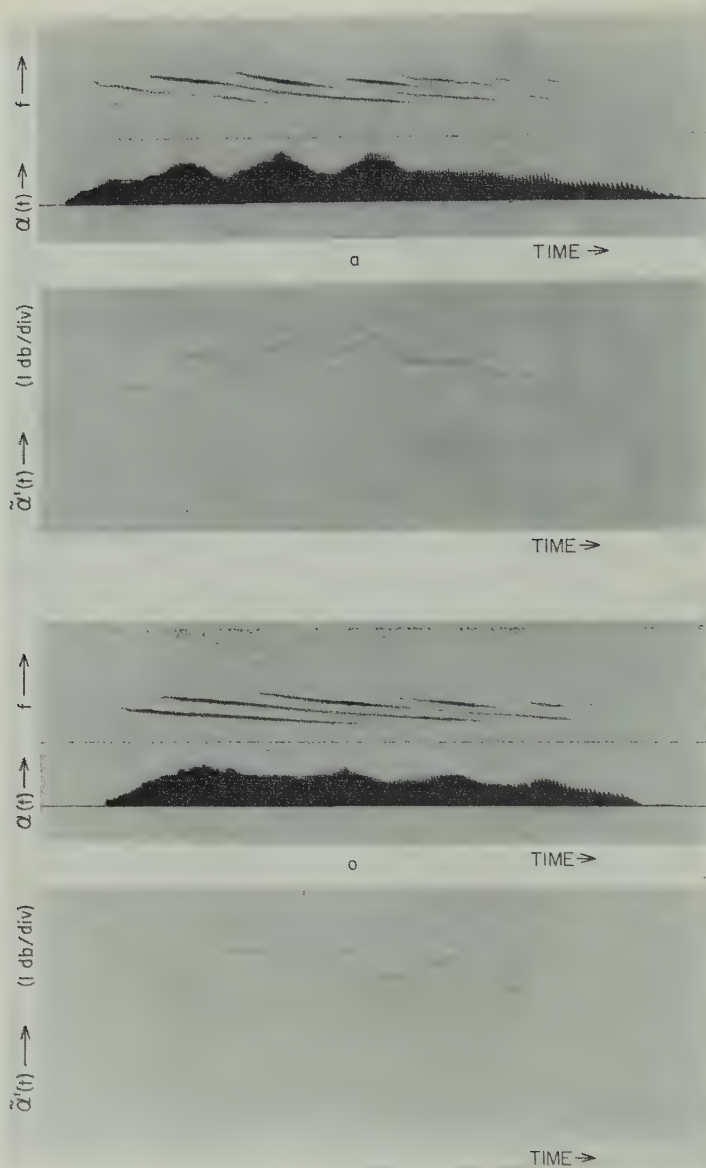


Fig. 8. Narrow-band spectrograms, and amplitude and speech power curves for the vowels [a] and [o] pronounced with falling fundamental frequency.

filter can be employed to provide an approximation to a Gaussian filter.



6. Compression of the filter output according to a logarithmic function is easy and is desirable.

7. It is necessary to use care in the interpretation of data from a speech power circuit. An example of the complexities is the variation in the power level of a vowel which occurs during changes in its fundamental frequency.

### *Summary*

The measurement of speech power is basic to research on the acoustic correlates of linguistic stress. Two units essential to speech power measurement are a rectifier and a filter. An amplitude compressor is also desirable.

Theoretically a square-law rectifier is ideally suited for speech power measurements, but there are considerable practical difficulties in implementing such a circuit. The common "linear rectifier" has been found experimentally to be suitable for this application. A simple filter design may be achieved which is adequate for most applications and which avoids the theoretical and practical difficulties of several other filter designs discussed. Logarithmic amplitude compression of the filter output is desirable and is easily accomplished with a diode.

In making speech power measurements with the circuit developed, an effect previously discussed by *House* has been observed.

### *Die sprachlich-akustische Kraftmessung*

#### *Zusammenfassung*

Die Erforschung der physikalischen Korrelate des dynamischen Akzentes setzt eine Messung der akustischen Kraft beim Sprechen voraus. Zwei wesentliche Bestandteile eines Meßapparates sind ein Gleichrichter und ein Filter. Ein Amplitudenkompressor ist ebenfalls wünschenswert.

Theoretisch ist ein quadratischer Gleichrichter für die Messung der Sprechkraft ideal geeignet; bedeutende praktische Schwierigkeiten ergeben sich aber bei der Verwirklichung einer solchen Schaltung. Der gewöhnliche lineare Gleichrichter hat sich erfahrungsgemäß für diese Verwendung als passend erwiesen. Ein einfacher Filter ist konstruiert worden, welcher für die meisten Anwendungen genügt und welcher die theoretischen und praktischen Schwierigkeiten anderer hier besprochener Filtertypen vermeidet. Die erwünschte logarithmische Amplitudenkompression des Filterausgangssignals läßt sich mittels einer Diode leicht bewerkstelligen.

Während einiger Messungen der Sprechkraft mittels der hier beschriebenen Anordnung ist ein zuerst von *House* besprochener Effekt beobachtet worden.

### *La mesure de l'énergie de la parole*

#### *Résumé*

La mesure de l'énergie de la parole est fondamentale dans les recherches sur les corrélatifs acoustiques de l'accent. Deux appareils sont essentiels pour la mesure de

cette énergie: un rectificateur et un filtre. Il est aussi souhaitable d'avoir un compresseur d'amplitudes.

En théorie, un rectificateur en règle de carrés convient idéalement à la mesure de l'énergie de la parole, mais on rencontre de considérables difficultés pratiques quand il s'agit de réaliser un tel circuit. Le rectificateur linéaire ordinaire s'est montré, à l'expérience, apte à cet emploi. On peut réaliser une simple forme de filtre qui satisfait à la plupart des emplois et qui évite les difficultés théoriques et pratiques de plusieurs autres formes de filtres décrites ici. La compression d'amplitude logarithmique du filtrage est souhaitable et facilement réalisable grâce à une diode.

En mesurant l'énergie de la parole au moyen du circuit ci-dessus, on a observé un effet étudié auparavant par *House*.

### References

1. *Cooper, F. S.; Liberman, A. M. and Borst, J. M.*: The interconversion of audible and visible patterns as a basis for research in the perception of speech. *Proc. nat. Acad. Sci., Wash.* 37: 318-325 (1951).
2. *Fant, G.*: Acoustic analysis and synthesis of speech with applications to Swedish. *Ericsson Technics no. 1*: 3-108 (1959).
3. *Fant, G.*: Modern instruments and methods for acoustic studies of speech. *Proc. 8th int. Congr. Ling.*, pp. 282-358 (Oslo 1958).
4. *Fletcher, H.*: Speech and hearing in communication (New York 1953).
5. *Fry, D. B.*: Experiments in the perception of stress. *Language and Speech 1*: 126-152 (1958).
6. *House, A. S.*: A note on optimal vocal frequency. *JSHR 2*: 55-60 (1959).
7. *Klapper, J. and Harris, C. M.*: On the response and approximation of Gaussian filters. *IRE Transactions on Audio AU-7*: 80-87 (1959).
8. *Lehiste, I.*: Some acoustic correlates of accent in Serbo-Croatian. (To appear in *Phonetica*.)
9. *Lehiste, I. and Peterson, G. E.*: Vowel amplitude and phonemic stress in American English. *J. acoust. Soc. Amer.* 31: 428-435 (1959).
10. *Licklider, J. C. R.*: Basic correlates of the auditory stimulus. *Handbook of Experimental Psychology*, pp. 985-1039 (John Wiley and Sons, Inc., New York 1951).
11. *Stevens, S. S.*: Calculation of the loudness of complex noise. *J. acoust. Soc. Amer.* 28: 807-832 (1956).
12. *Voss, H. H.*: Realisierbare Tiefpässe und Bandpässe minimaler Phase mit geebneten Laufzeit und aperiodischem Einschwingverhalten. *Frequenz 8*: 98-102 (1954).

Author's address: Prof. Dr. Gordon E. Peterson, The University of Michigan, Communication Sciences Laboratory, Ann Arbor, Mich. (USA)



## Phonèmes labio-vélaires en espagnol ?

Par BERTIL MALMBERG, Lund

On a beaucoup discuté l'interprétation phonologique de complexes phonétiques tels que les diphtongues, les affriquées et certaines consonnes à «double localisation» (palatalisées, labio-vélarisées, etc.). Je me dispense de renvoyer ici à toutes les contributions qui ont été données à la solution de ce problème depuis la fondation de l'Ecole de Prague – et même avant, surtout par le grand phonéticien auquel est dédié ce volume. Je ne fais que signaler ici le principe généralement valable pour l'interprétation phonologique, ou phonémique, de cas pareils. Si, dans un système donné, il existe par exemple une suite phonétique  $[tʃ]$  et aussi, admissibles dans la même position, un  $[t]$  et un  $[ʃ]$  – donc les trois possibilités  $[tA]$ ,  $[ʃA]$  et  $[tʃA]$  – on interprète  $[tʃ]$  comme un groupe de phonèmes («cluster»). Dans ce cas,  $|tʃ/|$  s'oppose à  $|tr/|$ ,  $|ts/|$ , etc. Si, par contre, le système connaît  $[t]$  et  $[tʃ]$  mais pas de  $[ʃ]$ , il devient nécessaire de voir dans  $[tʃ]$  un seul phonème  $|tʃ/|$ , opposé à  $|t/|$ , etc. Un segment phonétique qui est conditionné par son entourage et qui, par conséquent, suppose celui-ci, ne peut pas être un phonème, tout au plus le trait distinctif d'un autre phonème ( $[ʃ]$  étant dans notre deuxième exemple le trait distinctif qui oppose  $|tʃ/|$  à  $|t/|$ ). C'est là grosso modo le principe, et ce principe me semble bon. Reste à voir pourtant si, dans certains cas et dans tel ou tel but spécifique, il ne pourrait pas être légitime d'essayer une autre interprétation qui, dans ce cas, serait justifiée par le résultat auquel elle aboutit.

L'espagnol connaît, comme plusieurs des langues sœurs, les groupes (orthographiques) *cu-*, *gu-*, phonétiquement  $[kw]$ ,  $[gw]$ , dans par exemple *cuándo*, *cuál*, *guante*, *guardar* (celui-là remontant historiquement à un groupe latin correspondant, écrit *qu-*, dont nous dispensons de discuter ici le caractère phonologique; celui-ci représentant le plus souvent un *w-* germanique, ou d'origine amérindienne ou autre). La langue connaît aussi un  $|k/|$  et un  $|w/|$ , ce

dernier écrit *hu-* (*huevo*, *hueso*, etc.) en orthographe ordinaire<sup>1</sup>. Il me semble donc légitime, en vertu du principe d'interprétation que nous venons de formuler, d'interpréter les groupes */kw/* et */gw/* comme biphonématiques, puisque le */w/* initial est possible devant voyelle (quoiqu'avec une distribution limitée)<sup>2</sup> et que le */k/* dans la même position n'est pas automatiquement suivi de */w/*. De même pour */gw/*.

Dans mon article «Occlusion et spirance dans le système phonologique de l'espagnol» (*Mélanges Michaëlsson*, 1952, pp. 356-365), j'ai dressé une liste des dites occlusives de la langue avec leurs variantes faibles (spirantes), donc dans la série des sonores: *b - b̃*, *d - ð*, *g - γ*. J'y ai indiqué aussi les réductions qui, surtout dans les formes dialectales et vulgaires de la langue, se produisent en fin de syllabe où très peu des distinctions admises à l'initiale restent valables. Je renvoie pour tous les détails à cet article et me contente de rappeler ici que, dans ce schéma, je me suis permis de compter les groupes *cu-* */kw/* et *gu-* */gw/* (variante faible *[γw]*) comme des phonèmes à part, opposés comme des labio-vélaires aux */k/* et */g/* non labiaux, donc en transcription phonétique *[kʷ]* et *[gʷ]* (*[γʷ]*), phonologiquement */kʷ/* et */gʷ/*.

Une conséquence de cette interprétation est que le *fu-* de *fuego*, *fuera* sera analysé comme */fʷ/*, et le *ju-* de *juego*, *jueves* comme */xʷ/*, opposés à */f/* et à */x/* «simples». C'est là du reste une opposition à rendement très faible, *fuego* ∼ *juego* étant la seule paire minimale.

Il sera donc possible de dresser le schéma suivant des consonnes

<sup>1</sup> Je pars, en faisant l'analyse ainsi, de la description donnée par Tomás Navarro, *Pronunciación española*, § 65 (et la note 1, p. 64), selon lequel la langue vulgaire seule peut aller jusqu'à une occlusive dans ces mots (*huevo*, *hueso*), après nasale et en position initiale absolue, tandis que déjà la conversation familière peut développer une véritable consonne spirante (un *[w]*) dans *ahuecar* [*awe'kar*], *hueso* [*'wesɔ*], etc. Déjà l'orthographe utilisée pour rendre dans la littérature en langue populaire la prononciation *[gʷ]* ou *[γʷ]*, à savoir *gü-* pour *hu-*, prouve que cette différence de traitement entre la prononciation châtiée et la prononciation populaire est sentie consciemment par les personnes qui se servent de ces graphies et que, par conséquent, l'initiale de *guante* est à regarder comme différente de celle de *hueso*. (Je reviens plus loin à la description différente de ces faits que donne Emilio Alarcos Llorach, *Fonología española*, 3 éd., 1961.) Sinon, il n'y aurait aucune raison de modifier l'orthographe dans les textes en langue vulgaire.

<sup>2</sup> Si l'on fait abstraction du verbe *huir* et dérivés (*huída*) et de quelques américanismes avec des doublets en *gu-* (*huíncha - güíncha*, *huira - güira*), */w/* (orth. *hu-*) ne se retrouve que devant *-e* (*hueco*, *huerto*, *hueso*, etc.). Cf. aussi plus loin. Les faits historiques qui expliquent cette distribution sont trop évidents pour que j'insiste. L'alternance morphologique entre */w/* consonne et des formes avec une voyelle dans le radical dans *huir* ∼ *huye*, etc., *huele* ∼ *oler*, etc. est une complication des faits phonologiques dont je fais abstraction pour le moment.

(occlusives et spirantes) qui nous intéressent (en excluant pour des raisons de simplification systématique les nasales, les liquides et la sifflante *s*) avec leurs principales variantes:

$$\begin{array}{rcl}
 p & - & b \text{ (} b \text{)} - f \text{ (} \varphi \text{)}^3 \\
 t & - & d \text{ (} \delta \text{)} - \text{p} \\
 t^{\wedge} \text{f} & - & d^{\wedge} \text{j} \text{ (} j \text{)}^4 \\
 k & - & g \text{ (} \gamma \text{)} - x \\
 k^w & - & g^w \text{ (} \gamma^w, w \text{)} - x^w \\
 & & \begin{array}{ccc} | & & | \\ w & & f^w \text{ (} \varphi^w \text{)}^3 \end{array}
 \end{array}$$

Il ressort de ce schéma qu'il y a quelques anomalies dans le système. Les spirantes sonores sont normalement les variantes faibles de phonèmes neutres au point de vue de la distinction occlusion  $\infty$  spirance, selon les règles connues (*Valencia* [ba'lenɲja] mais *a Valencia* [a ba'lenɲja]). La différence entre [g<sup>w</sup>] et [w] au contraire a l'air de pouvoir être phonologique (même si les paires minimales font défaut à cause de la distribution asymétrique des phonèmes). Ce serait donc le dernier reste d'une opposition que le castillan a perdue au cours du moyen âge et que les efforts des maîtres d'école (surtout, semble-t-il, dans certains pays de l'Amérique latine; cf. mes *Etudes* précitées, pp. 61-62, et mes *Notas sobre la fonética del español en el Paraguay*, 1947, pp. 15-16) n'ont pas réussi à rétablir (sauf dans la bouche de quelques pédants). Il n'y a donc rien de surprenant dans le fait bien connu que, dans un très grand nombre de parlers castillans, en Espagne aussi bien qu'en Amérique, cette distinction disparaît également et que les mots à *hu-* initial sont englobés dans le système d'alternance entre occlusive et spirante connu des autres séries, d'où *güeso*, *güevo*, etc. dans la littérature en langue populaire. La forme faible [w], réalisée à l'intervocalique dans *yo guardo* ['ɣ<sup>w</sup>ardo] ou ['warðo], est le point de départ naturel de

<sup>3</sup> La prononciation bilabiale du *f* est sans doute plus répandue dans les différents parlers castillans que ne le font supposer les indications des manuels et des descriptions dialectologiques. C'est surtout devant *u* que le *f* bilabial est fréquent, donc aussi dans le phonème [f<sup>w</sup>] de notre schéma (cf. mes *Etudes sur la phonétique de l'espagnol parlé en Argentine*, 1950, pp. 96-97).

<sup>4</sup> Il serait possible de classer le [s] comme la palatale spirante (sourde), parallèle aux [f] - [p] - [x]. On connaît le caractère apico-alvéolaire du [s] castillan, d'où un timbre assez «palatal». J'hésite pourtant un peu à le faire, vu le caractère phonétique particulier de ce phonème, mais vois, d'autre part, dans sa place un peu en dehors du reste du système une des raisons de sa tendance à devenir [h] en position implosive (cf. plus loin).



cette confusion. La différence entre  $[\gamma^w]$  et  $[w]$  est graphique plutôt que phonétique. Elle se réduit tout au plus à une minime différence de rétrécissement dorso-vélaire et n'est pas plus sensible que celle entre les différentes réalisations possibles du  $[\delta]$  ou du  $[\tilde{\delta}]$  intervocaliques, où toutes les nuances entre une véritable spirante à friction audible et un rétrécissement peu perceptible du passage de l'air sont imaginables et instrumentalement vérifiées (cf. mes *Etudes* précitées, pp. 63 ss, 70 ss, 85 ss). Dans bien des parlers castillans, le  $[w]$  a donc disparu complètement de l'inventaire des phonèmes consonantiques et avec celui-là aussi toute trace d'opposition entre occlusive et spirante sonores.

On sait aussi que la distinction entre *gu-* et *bu-* a été supprimée dans les mêmes parlers (*güeno* pour *bueno*, etc., et même quelquefois le phénomène inverse). C'est évidemment l'identité en position faible qui en est le point de départ,  $[w]$  étant la réalisation phonétique des deux: (*está*) *bueno*  $['weno]$ , (*compra*) *huevos*  $['weβos]$ .  $[w]$  peut donc facilement être senti comme la variante faible de  $[g^w]$ , d'où la réalisation de la forme forte de  $['weno]$  comme  $['g^weno]$  (orth. *güeno* dans la littérature en langue populaire). Il n'y a donc aucune raison de prendre en considération aussi un phonème vélarisé  $[b^w]$  opposé à  $[g^w]$ , puisque le type disparaît avec la stabilisation d'une série labio-vélaire à valeur phonologique.

Il serait sans doute justifié de prendre en considération ici également une possibilité d'interprétation des *gu-*  $\sim$  *hu-* qui est en partie différente de celle proposée ci-dessus. En ne tenant compte que de la position initiale, il serait tentant de voir dans  $[g^w]$  –  $[w]$  des variantes en distribution complémentaire, *gu-*  $[g^w]$  se trouvant devant  $[a]$ , *hu-*  $[w]$  devant  $[e]$ . Il n'y a pas d'exemples devant  $[o]$  et  $[u]$ ; ils sont rares devant  $[i]$ . Les exemples de *hu-* initial devant  $[a]$  dans le grand dictionnaire espagnol-danois de Carl Bratli (1947) sont tous des mots américains avec un doublet en *gu-* (*huaca* – *guaca*, *huango* – *guango*, etc.). Les quelques exemples de *gu-* devant  $[e]$  sont tous des mots spéciaux (américanismes, dialectalismes ou emprunts: *güecho*, *güelfo*). Devant  $[i]$ , il y a quelques exemples comme *güilo*, *güira*, *güica* (dont quelques-uns avec une variante en *hu-*: *huira*). A l'intervocalique, on retrouve les mêmes variations, surtout dans les américanismes: *ahuanés* – *aguanés*, etc. Mais il serait peut-être malgré tout aller trop loin que de voir dans ces deux initiales uniquement des variantes conditionnées du même phonème. Il y a tout de même une possibilité d'opposition, me semble-t-il, à l'in-

tervocalique entre *agüero* /a'g<sup>w</sup>ero/ et *ahuero* (du verbe *ahuerar*) /a'wero/. On trouvera facilement d'autres exemples avec *-gu-* et *-hu-* intervocaliques où, avec notre point de départ, il serait faux de parler d'identité phonétique. Mais de l'autre côté, il est évident que cette distribution est un des facteurs qui contribuent à la disparition définitive de toute trace de distinction. C'est cette distribution et la tendance structurale à la disparition de l'opposition qui expliquent ensemble le fait bien connu que tous les emprunts aux langues américaines à *hu-* [w] initial ont été hispanisés avec *gu-* (*guanaco*, *guano*, *guarango*, etc.; cf. mes *Etudes*, p. 88, *Mélanges Michaëlsson*, p. 359, *Tiscornia*, Bibl. dial. hispanoamericana III, p. 54). Rappelons aussi que le même renforcement d'une initiale en [we] a eu lieu dans d'autres dialectes hispaniques que le castillan (aragonais, aranais, etc.; des exemples aragonais chez *Tilander*, *Los furos de Aragón*, 1937, pp. XL-XLI). Donc, s'il était faux<sup>5</sup> de dire tout court qu'il n'y a pas de distinction du tout et que les deux types sont en distribution complémentaire, il serait également faux de parler tout simplement d'opposition. Nous nous trouvons en réalité devant une distinction en train de disparaître et devant un système en pleine reconstruction.

Dans notre schéma ci-dessus, le /f<sup>w</sup>/ est isolé. Son caractère labial est en réalité redondant sauf dans le cas de la paire unique mentionnée *fuego* ∼ *juego*. Par conséquent, cette caractéristique redondante disparaît avec, comme résultat, une confusion complète entre /f<sup>w</sup>/ et /x<sup>w</sup>/, d'où les formes *juí*, *jueron*, *juera*, etc. (phon. ['x<sup>w</sup>era], etc.) dans la littérature en langue populaire. Il est évident que cette confusion a lieu d'autant plus facilement que, justement dans cette combinaison, l'articulation est très souvent bilabiale [φ<sup>w</sup>] (cf. mes *Etudes*, pp. 96-97, de même le schéma ci-dessus, la note 3)<sup>6</sup>.

Dans les parlers qui connaissent les réductions phonologiques dont nous venons de donner des exemples, le /b/ du tableau ci-dessus manque en général lui aussi, confondu avec /s/. C'est, on le sait, le cas de toute l'Amérique espagnole et d'une grande partie du Sud de

<sup>5</sup> Toujours avec la description de *Tomás Navarro* comme base de la discussion.

<sup>6</sup> C'est par un raisonnement en principe analogue que j'ai cru pouvoir expliquer le fameux passage *f* > *h* en castillan pré littéraire (la perte d'un trait redondant; voir mon article dans les *Mélanges Petrovici*, sous presse). A mon avis, ce passage devient également plus naturel si l'on part, comme je pense qu'il le faut, d'un *f* bilabial.

la Péninsule<sup>7</sup>. Dans ce cas, c'est le /s/ qui prend la place du /b/ dans notre schéma, partenaire spirant sourd de /t/ et de /d/, ce qui implique à son tour que le caractère sifflant de cette spirante («strident» opposé à «mellow», dans la terminologie de Jakobson, Fant et Halle, *Preliminaries*, § 2.321) devient redondant. Le système ne permet plus d'opposer une spirante à ouverture «ronde» à une autre à ouverture «plate». Il me semble donc parfaitement normal que ce soit dans ces mêmes parlers (sans /b/) que le /s/ tende surtout à s'affaiblir, à devenir une espèce de [h], spirante neutre se réglant d'après l'entourage<sup>8</sup>. Mais déjà la structure du système castillan optimal (avec /b/) est favorable à un tel développement (cf. ci-dessus, note 4). Je renvoie pour des exemples de tous ces phénomènes à mes *Etudes* précitées (pp. 156-172) et aux nombreux renvois qui y ont été faits. Cet affaiblissement a lieu, on le sait, d'abord et surtout en fin de syllabe, en vertu de la tendance syllabique de la langue (dont j'ai parlé ailleurs; cf. Boletim de filologia IX, 1948, pp. 99-120<sup>9</sup>), mais s'étend, dans quelques parlers, au /s/ intervocalique (intérieur de l'Argentine, Chili, Venezuela, etc. en Amérique; *Etudes*, p. 168), et même quelquefois initial (Lenz, *Bibl. dial. hispanoamericana* VI, p. 125). Il est normal que la redondance ait plus de peine à se maintenir en position faible (implosive) qu'en position forte (initiale de la syllabe).

Ces quelques remarques ont été faites pour illustrer une question de principe et de méthode plutôt que pour contribuer à l'analyse phonémique du castillan et de ses tendances évolutives, sur lesquelles j'aurai du reste bientôt l'occasion de revenir. Je me suis permis, dans mon analyse, de partir de l'hypothèse d'une interprétation monophonématique des *cu-*, *gu-*, *fu-*, *ju-* castillans (ou plus correctement des groupes ainsi orthographiés) – interprétation

<sup>7</sup> Voir pour des précisions mes *Etudes* précitées, pp. 172-173, et les renvois y faits (Alonso-Lida, *Bibl. dial. hispanoamericana* VI, p. 213, note, etc.).

<sup>8</sup> Il est probable que le caractère prédorsal du *s* dans ces régions facilite l'affaiblissement. On sait que le *s* castillan du centre est apical. Ce n'est certainement pas non plus un effet du hasard si c'est en principe dans les régions à opposition /s/ ∼ /b/ que le *s* est apical et si ce sont les parlers à une seule sifflante qui préfèrent le *s* apical. Il importe de rappeler que l'espagnol ne connaît pas de [ʃ]. Je note enfin aussi la possibilité d'une tendance, au moins en argentin populaire, à confondre -s (en fin de syllabe) avec -f, d'où la forme (graphique) *astosa* pour [fièvre] *afiosa* 'fièvre aphteuse', confusion faite probablement à partir de [a<sup>h</sup>tosa] avec «aspiration» du *s* et de [aφ'tosa] respectivement, différence extrêmement subtile auditivement (voir mes *Etudes*, p. 97).

<sup>9</sup> C'est là une des questions sur lesquelles je pense revenir sous peu dans un autre ordre d'idées.



fausse d'après les principes généralement admis en analyse phonémique. Mais ce point de départ m'a permis de rendre compte – et de donner une explication structurale – d'une tendance populaire très forte et très répandue dans le domaine de langue castillane, tendance qui, du reste, entre d'une façon naturelle dans le cadre des modifications structurales subies par ce dialecte roman depuis l'époque pré littéraire, c'est-à-dire au cours du dernier millénaire. La même interprétation m'a permis aussi de mieux comprendre la disparition du  $/f^w/$ , confondu avec  $/x^w/$  pour une raison qui ressemble étrangement à l'énigmatique disparition du  $f$  du système castillan au début de l'ère littéraire<sup>10</sup>.

La dite aspiration du  $s$  espagnol – qui met son empreinte sur tant de parlers castillans, péninsulaires et américains – obtient aussi sa raison d'être à l'intérieur du système ainsi esquissé, quoique ce phénomène ne soit pas en rapport direct avec les autres et qu'ils ne couvrent pas non plus exactement les mêmes aires dialectales. Il est donc probable que le système réalisé en castillan vulgaire et dialectal est celui que la langue cherche à réaliser mais que les tendances conservatrices empêchent de se généraliser. C'est donc à travers les phénomènes dialectaux, et en appliquant comme hypothèse de travail une interprétation au fond inexacte, que je crois pouvoir décrire le système consonantique de l'espagnol tel qu'il se manifeste dans ses tendances évolutives et dans ses formes les plus évoluées (= les plus pauvres au point de vue du nombre des distinctions utilisées). Il est évident que le système évolue vers une confusion des groupes (orthographiés)  $hu-$  et  $gu-$ . Dans les états de langues nés de cette évolution, le  $[w]$  n'existe plus comme phonème, donc  $/g^w/$  doit s'interpréter comme un seul phonème. C'est en appliquant cette même interprétation déjà au système non évolué (le castillan officiel) qu'on arrive le mieux à expliquer les modifications subies par celui-ci. C'est donc le résultat de cette même évolution qui

<sup>10</sup> Il est intéressant de constater que la position devant  $/u/$  est la seule où le passage  $f > h$  n'ait pas eu lieu en castillan pré littéraire (abstraction faite des mots savants et des groupes avec liquides), empêché, d'après l'opinion généralement admise, par une différenciation conservante. Les tendances dialectales et vulgaires ont donc tout simplement généralisé le passage  $f > h$ , avec la seule différence, pourtant, que la spirante postérieure (pharyngale et laryngale respectivement, mais avec une tendance nette à la généralisation de l'articulation pharyngale aussi dans le dernier cas:  $[h^w]$ , si l'on me permet cette fantaisie, ou peut-être mieux  $[n]$ ) dans le cas de la labiovélaire n'a pas encore été amuie. Cf. pourtant une certaine tendance à confondre en argentin vulgaire aussi  $/g^w/$  et  $/x^w/$  (cf. mes *Etudes*, p. 97).

justifie l'hypothèse de travail choisie au début de l'analyse<sup>11</sup>. C'est cette hypothèse qui permet de comprendre comment et pourquoi les tendances structurales du système consonantique espagnol se généralisent et comment sa structure arrive peu à peu à un équilibre parfait, toutefois avec, comme conséquence, la perte de quelques-unes de ses distinctions les plus utiles.

En interprétant nos groupes comme une suite de deux phonèmes il aurait été plus difficile, ou en tout cas plus compliqué, de rendre compte de l'introduction du phonème /g/ devant le /w/ (ou /u/) initial. Les méthodes d'analyse, en linguistique comme partout ailleurs, doivent au moins en partie être jugées d'après les résultats qu'on obtient en les appliquant<sup>12</sup>.

<sup>11</sup> Dans sa *Fonología española* Alarcos Llorach objecte contre l'interprétation monophonématique des *cu-*, *gu-*, qu'avec cette analyse on serait obligé d'introduire toute une série d'autres labio-vélaires: /b<sup>w</sup>, p<sup>w</sup>, t<sup>w</sup>, d<sup>w</sup>, n<sup>w</sup>, l<sup>w</sup>/ (dans *vuelta, puerto, tuerto, duelo, nuevo, luego*, etc.). La remarque est parfaitement logique, mais tandis que l'interprétation monophonématique des groupes initiaux de *luego, nuevo* n'aurait aucune autre conséquence qu'un allongement inutile et méthodiquement inacceptable de la liste des consonnes, la même interprétation des /k<sup>w</sup>/ - /g<sup>w</sup>/ rendra plus clairs certains phénomènes d'évolution dialectale et vulgaire en les faisant entrer dans le cadre de tendances structurales d'un caractère général.

<sup>12</sup> Selon la description des faits discutés ici donnée par Emilio Alarcos Llorach (*Fonología española*, 3 éd. 1961, § 103, pp. 157-159), l'altération du système décrite ici serait déjà un fait accompli. Selon cet auteur, il n'y aurait aucune différence de prononciation entre *gu-* et *hu-* (les deux groupes orthographiques étant en distribution complémentaire; «pura cuestión ortográfica», p. 158, cf. ci-dessus, note 2). Il dit expressément que *de huellas* et *degüellas* se prononcent de façon identique («se realizan igual», p. 158). Y aurait-il entre les descriptions données par ces deux spécialistes espagnols une différence de génération? C'est possible. Je n'ose pas me prononcer là-dessus. Menéndez Pidal (*Manual*, § 35, 7e) exprime une opinion plutôt intermédiaire, en admettant d'un côté la possibilité d'une prononciation occlusive du type *huevo* mais en donnant d'autre part la transcription [ũwerto] ou [ũŋ'werto] à côté de [ũŋ'gwerto]. Peu importe. Si c'est la description donnée par M. Alarcos qui est la bonne, ceci prouve tout simplement que la tendance structurale en question a abouti et que l'anomalie qu'était l'existence d'un /w/ opposé à un /g<sup>w</sup>/ a disparu. On sait qu'en réalité la prononciation décrite ici comme populaire (*güeso, güerto*, etc.) ne date pas d'hier et que les graphies de ce genre abondent déjà dans la littérature classique. C'est un exemple parmi tant d'autres d'un vulgarisme castillan qui est en même temps un archaïsme. Donc, la tendance à équilibrer le système est ancienne. Est-ce grâce à l'orthographe et aux efforts des grammairiens que la généralisation définitive de l'équilibre phonologique n'a réussi qu'avec peine et à une époque toute récente à s'imposer dans la bonne prononciation espagnole? - M. Alarcos (*op. cit.*, p. 159, note 1) fait remarquer que M. Martinet (*Economie des changements phonétiques*, p. 83) accepte le /w/ parmi les phonèmes de la langue mais que Amado Alonso le trouve «muy objetable» (Bibl. dial. hispanoamericana VI, 1940, p. 88, note 1). Je fais remarquer pourtant qu'à l'endroit cité Alonso ne discute pas l'existence du /w/ dans le système consonantique, seulement le caractère monophonématique du groupe [ɣw], et que c'est cette interprétation qu'il trouve «muy objetable».

### Résumé

L'auteur propose comme hypothèse de travail une interprétation monophonématique des  $/k^w/$ ,  $/g^w/$  espagnols (orthographiquement *cu-*, *gu-*). Cette interprétation a l'avantage de mieux rendre compte d'une tendance évolutive, caractéristique du phonétisme castillan depuis le moyen âge, à savoir la confusion du phonème  $/w/$  (écrit *hu-*) avec  $/g^w/$ , et aussi de la confusion des groupes écrits orthographiquement *gu-* et *bu-* (*bueno* > *güeno*, etc.). Avec la confusion entre  $/w/$  et  $/g^w/$ , la paire labio-vélaire  $[g^w] - [w]$  s'est réglée d'après le même système d'alternance entre occlusive et spirante que, depuis la fin du moyen âge, les autres trois paires  $[b] - [\beta]$ ,  $[d] - [\delta]$  et  $[g] - [\gamma]$ .

### Zusammenfassung

Der Verfasser schlägt als Arbeitshypothese vor, spanische  $/k^w/$ ,  $/g^w/$  (orthographisch *cu-*, *gu-*) monophonematisch zu werten. Eine solche Wertung hat den Vorzug, daß sie besser erklärt, wie sich die spanischen Laute seit dem Mittelalter entwickelt haben. Die Phoneme  $/w/$  (geschrieben *hu-*) und  $/g^w/$  sind zusammengefallen. Außerdem sind die Gruppen, die *gu-* und *bu-* geschrieben werden, zusammengefallen (*bueno* > *güeno* usw.). Nach dem Zusammenfall von  $/w/$  und  $/g^w/$  nimmt das labio-velare Paar  $[g^w] - [w]$  an dem gleichen alternativen System von Verschluß- und Reibelaut teil wie die drei anderen Paare  $[b] - [\beta]$ ,  $[d] - [\delta]$  und  $[g] - [\gamma]$  seit dem Ende des Mittelalters.

### Summary

The author puts forward as a working hypothesis a monophonematic interpretation of the Spanish  $/k^w/$ ,  $/g^w/$ , (*cu-*, *gu-* in ordinary orthography). This interpretation has the advantage of giving a better explanation of a tendency in Spanish phonetics which has been at work since the Middle Ages, i.e. the confusion between the phonemes  $/w/$  (written *hu-*) and  $/g^w/$ , and also of the confusion between the groups written in ordinary orthography *gu-* and *bu-* (*bueno* > *güeno*, etc.). After the confusion between  $/w/$  and  $/g^w/$ , the labio-velar pair  $[g^w] - [w]$  has been adapted to the same system of alternation between stop and fricative as, since the end of the Middle Ages, characterizes the other three pairs  $[b] - [\beta]$ ,  $[d] - [\delta]$ , and  $[g] - [\gamma]$ .

### Bibliographie

- Alarcos Llorach, E.: *Fonología española*. 3 éd. (Madrid 1961).  
 Alonso, A.: *Problemas de dialectología hispanoamericana* (Bibl. dial. hispanoamericana I: Buenos Aires 1930).  
 Jakobson, R.; Fant, G. and Halle, M.: *Preliminaries to speech analysis* (MIT Acoust. Lab. Techn. Rep. 13, Cambridge 1952).  
 Lenz, R.: *El español en Chile*. Traducción, notas y apéndices de A. Alonso y R. Lida (Bibl. dial. hispanoamericana VI, Buenos Aires 1940).



- Malmberg, B.*: Notas sobre la fonética del español en el Paraguay (Vetenskaps-Societeten i Lund, årsbok 1947). – La structure syllabique de l'espagnol (Boletim de filologia IX, 1948). – Etudes sur la phonétique de l'espagnol parlé en Argentine (Etudes romanes de Lund X, Lund 1950). – Occlusion et spirance dans le système consonantique de l'espagnol (Mélanges Michaëlsson, Göteborg 1952). – Le passage  $f > h$  – perte d'un trait redondant? (Mélanges Petrovici, Cluj, sous presse).
- Martinet, A.*: Economie des changements phonétiques (Francke, Bern 1955).
- Menéndez Pidal, R.*: Manual de gramática histórica española, 7 éd. (Madrid 1944).
- Navarro, T.*: Manual de pronunciación española. 5 éd. (New York 1957).
- Tilander, G.*: Los fueros de Aragón (Acta Reg. Soc. Humaniorum Litt. Lundensis XXV, Lund 1937).
- Tiscornia, E. F.*: La lengua de «Martín Fierro» (Bibl. dial. hispanoamericana III, Buenos Aires 1930).

Adresse de l'auteur: Professeur Bertil Malmberg, Fonetiska institutionen, Kävlingevägen 20, Lund (Suède)

Physikalisch-Technische Bundesanstalt, Braunschweig und Berlin

## **Zur Technik der Tonbandaufnahme bei Sprachuntersuchungen**

Von WERNER KALLENBACH und HANS-JÜRGEN SCHROEDER

### *Einleitung*

Die Magnettontechnik hat sich in den letzten 25 Jahren zu dem hochwertigsten Schallaufzeichnungsverfahren entwickelt. Sie ist im Bereich des Rundfunks nicht zuletzt wegen ihrer betrieblichen Vorteile: ständige Aufnahmebereitschaft, sofortige Wiedergabemöglichkeit der Aufnahmen ohne mechanische oder chemische Nachbehandlung und einfache Möglichkeit des Schneidens und Klebens zu einem unentbehrlichen Bestandteil der Studioeinrichtungen geworden. Neben den hochwertigen Studiomaschinen wurden in den letzten Jahren von der Industrie in großem Maßstabe auch einfach zu bedienende Heimgeräte entwickelt, die wegen ihres verhältnismäßig niedrigen Preises einen weiten Abnehmerkreis fanden. Durch die Herabsetzung der Bandgeschwindigkeit, die Einführung der Zweispur- und neuerdings der Vierspurtechnik gelang es ferner, die Kosten für das Bandmaterial zu senken. Entsprechend den gesteigerten Anforderungen wurden auch neue Bandtypen entwickelt; trotzdem war ein gewisser Verlust an Aufzeichnungsqualität bei diesen Geräten nicht zu umgehen, doch ist dies bei dem in Frage kommenden Käuferkreis meist nicht von entscheidender Bedeutung.

Auch in der Wissenschaft und Forschung erkannte man bald die Vorteile der magnetischen Schallaufzeichnung und machte sie sich zunutze. In der Akustik werden beispielsweise die Schallvorgänge auf Tonband aufgenommen und die Aufzeichnungen anschließend im Laboratorium nach den verschiedensten Methoden untersucht.

Ebenso werden in der Phonetik, bei Dialektuntersuchungen,

aber auch in der Medizin bei Untersuchungen an Sprachgestörten und Geisteskranken sowie bei psychologischen und physiologischen Fragestellungen Tonbandaufnahmen in zunehmendem Maße verwendet. In allen diesen Fällen tritt an den Versuchsleiter das Problem heran, das für einen bestimmten Zweck geeignete Gerät zu beschaffen und insbesondere zu entscheiden, ob ein einfaches Heimgerät ausreicht oder eine hochwertige und dementsprechend kostspielige Studioapparatur erforderlich ist. Diese Frage läßt sich allgemein nicht beantworten, sie hängt von dem angestrebten Ziel der Untersuchungen ab und muß von Fall zu Fall geprüft werden.

Die folgenden Ausführungen sollen dazu beitragen, dem Versuchsleiter die Entscheidung zu erleichtern und ihm die Qualitätseinbuße bei kleinen Bandgeschwindigkeiten und ihre Auswirkungen auf die verschiedenen Sprachuntersuchungs-Methoden vor Augen zu führen. Dabei soll auch auf einige Punkte hingewiesen werden, die bei der Aufnahme selbst von Bedeutung sind.

### *Pegelschwankungen*

Die Herabsetzung der bei Studiogeräten üblichen Bandgeschwindigkeit von 38,1 cm/s auf 19,05, 9,5 oder sogar 4,75 cm/s bei Heimgeräten hat zur Folge, daß die auf dem Bände aufgezeichnete Information entsprechend im Verhältnis 1:2, 1:4 oder gar 1:8 zusammengedrängt wird, d. h. daß die auf dem Bände aufgezeichneten Wellenlängen entsprechend kürzer werden. Im gleichen Maße wie sich die Wellenlänge verkleinert, tritt auch der Streufluß zwischen benachbarten magnetischen Nord- und Südpolen in der Schicht weniger weit aus der Bandoberfläche hervor. Um den gesamten magnetischen Streufluß vom Wiedergabekopf erfassen zu können, muß das Band ständig mit ausreichendem mechanischen Kontakt am Kopf vorbeigeführt werden. Dieser Kontakt läßt sich in der Praxis jedoch nicht absolut sicher aufrechterhalten. Abhebungen von einigen  $\mu\text{m}$ , verursacht durch Unebenheiten der Bandoberfläche und aufliegende Staubpartikel lassen sich nicht vermeiden. Während solche Abhebungen bei den großen Bandgeschwindigkeiten keine Wirkung zeigen, führen sie bei kleinen Geschwindigkeiten, ebenso wie unvermeidliche Inhomogenitäten der magnetischen Schicht, zu erheblichen Pegelschwankungen. Zur Veranschaulichung dieses Effektes sind in Abb. 1 Pegelaufzeichnungen von 20 sec Dauer der Frequenzen 60 Hz, 1, 6 und 10 kHz wiedergegeben, die sich bei 38,



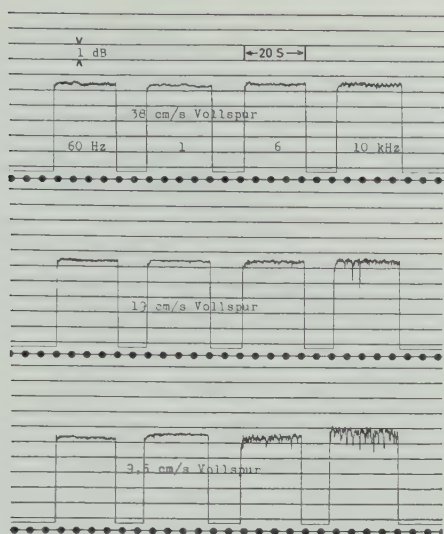


Abb. 1. Pegelschwankungen bei Vollspuraufzeichnung.

19 und 9,5 cm/s Vollspuraufzeichnung ergaben. Während bis 1 kHz bei allen Geschwindigkeiten die Schwankungen innerhalb eines Bereiches von 0,2 dB liegen, erreichen sie bei 10 kHz und 9,5 cm/s bereits Werte von 1,5 dB. Noch größer werden die Schwankungen, wie Abb.2 erkennen läßt, wenn man zur Halbspurabtastung übergeht, da eine kleine lokale Unebenheit oder ein kleiner Fehler in der Schicht sich hier stärker auf den Pegel auswirken als bei Vollspur. Die beiden unteren Aufzeichnungen von Abb.2 ergaben sich mit einem Kombikopf, bei dem für Aufnahme und Wiedergabe der gleiche Kopf benützt wird. Dies ist an sich der günstigste Fall, da kleine Abweichungen der Spaltform bei Aufnahme und Wiedergabe in gleicher Weise eingehen und daher bei der Wiedergabe nicht störend in Erscheinung treten. Trotzdem werden bei 9,5 cm/s und 10 kHz Werte von 1,5 dB und bei 4,75 cm/s und 6 kHz sogar von fast 3 dB erreicht. Noch wesentlich unangenehmer sind die Verhältnisse bei der Vierspurtechnik in Verbindung mit der Bandgeschwindigkeit 4,75 cm/s. Hier bilden die «drop-outs» ein zurzeit von der Technik noch nicht ausreichend beherrschtes Problem.

### *Spaltstellung*

Noch schwerwiegender als die Pegelschwankungen infolge mangelhaften Bandkontaktes ist der Einfluß der Schiefstellung des

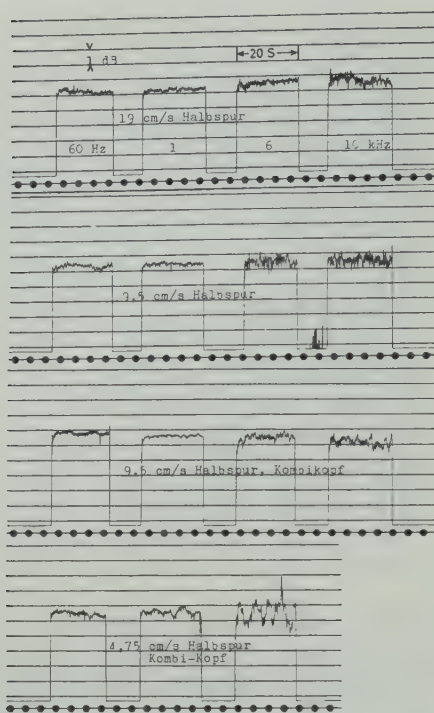


Abb. 2. Pegelschwankungen bei Halbspuraufzeichnung.

Spaltes bei kleinen Bandgeschwindigkeiten. Hierauf hat kürzlich *H. Schmidt*\* besonders hingewiesen. Verwendet man für Aufnahme und Wiedergabe das gleiche, mit einem kombinierten Sprech-Hör-Kopf ausgerüstete Gerät, so spielt es keine Rolle, ob der Kopfspalt genau senkrecht zur Längsachse des Bandes steht, da ja für die Abtastung der gleiche Kopf wie bei der Aufnahme verwendet wird und damit die Parallelität von Kopfspalt und Magnetisierungsrichtung innerhalb der Genauigkeit der Bandführung gewährleistet ist. Anders liegen die Dinge jedoch, wenn eine Aufnahme auf einem zweiten Gerät abgespielt werden muß. Dieser Fall kommt in der Praxis häufig vor, z.B., wenn Bandaufnahmen in einem anderen Institut ausgewertet werden sollen. Steht bei dem zweiten Gerät der Spalt nicht genau parallel zur Magnetisierungsrichtung des bespielten Bandes, so tastet die Oberkante des Spaltes eine gegenüber der Unterkante versetzte Stelle des Bandes ab. Es wird damit im

\* *H. Schmidt*: Die Justierung von Magnettonköpfen. Funktechnik 1960, S. 784–786.

Kopf eine Spannung induziert, die von der Oberkante bis zur Unterkante des Spaltes aus verschiedenen, in der Phase gegeneinander verschobenen Anteilen besteht. Im Extremfalle, wenn nämlich die Schiefstellung sich gerade über eine volle Wellenlänge der aufgezzeichneten Frequenz erstreckt, tritt bei der Wiedergabe sogar völlige Auslöschung ein. Bedenkt man, daß z.B. bei 9,5 cm/s und 10 kHz die Wellenlänge  $\lambda = c/v = 9,5/10000 \text{ cm} = 9,5 \text{ } \mu\text{m}$  beträgt, so erkennt man, daß schon Schiefstellungen von wenigen Winkelminuten einen erheblichen Verlust bei den höheren Frequenzanteilen zur Folge haben. *Schmidt* hat in seiner Arbeit den Verlust formelmäßig in Abhängigkeit von Bandgeschwindigkeit und Frequenz angegeben. Um zu zeigen, wie sich der Effekt in der Praxis auswirkt, sind in Abb. 3 die Frequenzgänge im Bereich von 1 bis 15 kHz dargestellt, die sich bei richtiger und um 6' abweichender Spaltstellung ergeben. Eine Abweichung von 6' ist nur ein geringer Betrag, wenn man bedenkt, daß die zur Einstellung von Tonbandgeräten benutzten Bezugsbänder, die einen besonderen Teil zur Spaltjustage besitzen, selbst schon Abweichungen von  $\pm 3'$  von der Senkrechtlage aufweisen können. In Abb. 3 sind von oben nach unten paarweise die Frequenzgänge für 38, 19 und 9,5 cm/s dargestellt und zwar links jeweils für Vollspur- und rechts für Halbspurabtastung. Zu beachten ist dabei, daß der Abstand zwischen zwei Ordinatenlinien hier 2,5 dB gegenüber 1 dB in Abb. 1 und 2 beträgt. Es ist deutlich zu erkennen, wie stark sich die Schiefstellung des Spaltes auf den Frequenzgang auswirkt, vor allem bei Vollspur. Außer dem Abfall der Höhen treten noch starke Pegelschwankungen infolge von Ungenauigkeiten der Bandführung auf. Bei Halbspurabtastung ist der Effekt wesentlich kleiner, weil der Abstand zwischen Spalt-Ober- und -Unterkante nur knapp halb so groß und damit die Versetzung in Längsrichtung des Bandes entsprechend geringer ist als bei Vollspur. Da bei Halbspur jedoch größere Pegelschwankungen und eine Verschlechterung des Rauschabstandes um mindestens 3 dB eintreten, bedeutet der Übergang zur Halbspuraufzeichnung immer eine Minderung der Qualität.

Ein Abfall des Frequenzganges bei hohen Frequenzen wirkt sich entsprechend auf Frequenzanalysen von Schallvorgängen aus und ist auch bei der Wiedergabe der Aufnahmen deutlich hörbar. Diesem Punkt ist daher bei Austausch von Tonbändern besondere Beachtung zu schenken. Bei Heimgeräten schwankt die vom Hersteller vorgenommene Einstellung von Gerät zu Gerät im allge-



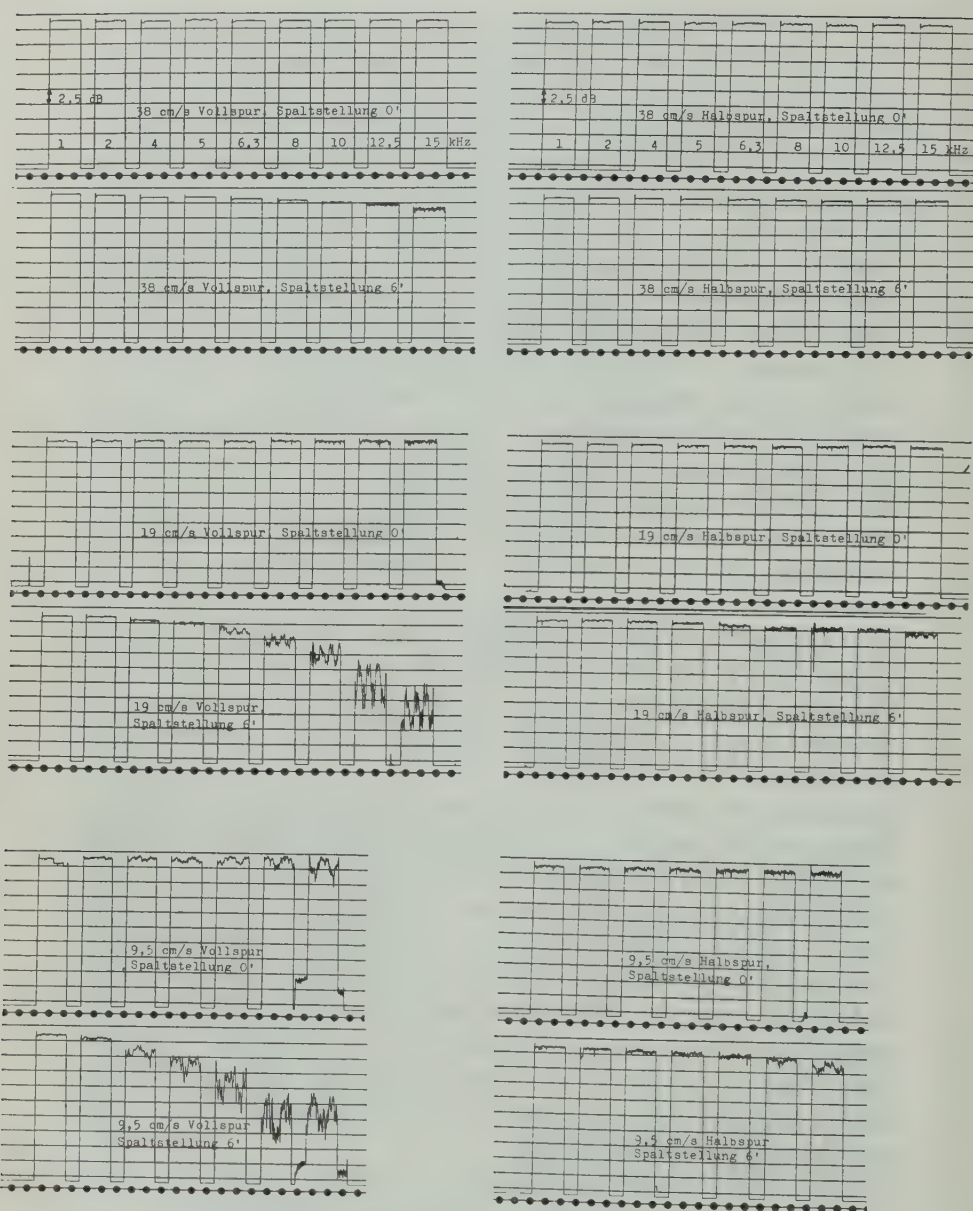


Abb. 3. Einfluß der Spaltschiefstellung auf den Wiedergabefrequenzgang.

meinen so stark, daß z. B. bei 9,5 cm/s Halbspuraufzeichnung in den meisten Fällen mit einem erheblichen Verlust an Höhen bei Austausch von Aufnahmen zu rechnen ist.

*Einfluß von Brummspannungen, Rauschen und Nachhall*

Es sollen nun einige Einflüsse betrachtet werden, die besonders schädlich sind, wenn von den Bandaufnahmen Aufzeichnungen der Tonhöhe mit dem Tonhöhenschreiber nach *Grützmacher und Lottermoser*\* vorgenommen werden sollen, was für eine quantitative Auswertung von Sprachaufnahmen häufig wünschenswert ist. Abb.4

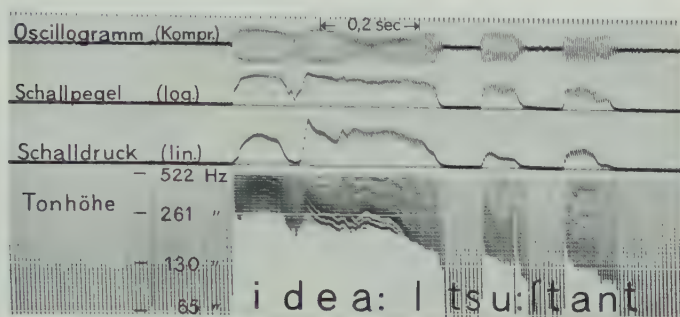


Abb. 4. Tonhöhenaufzeichnung des Wortes «Idealzustand».

zeigt als Beispiel die Aufzeichnung des Wortes «Idealzustand» in der Form, wie sie das Gerät jetzt liefert. Man erkennt in der Reihenfolge von oben nach unten die Zeitmarke (50 Hz), das Oscillogramm in komprimierter Form, um auch kleine Amplituden noch deutlich hervortreten zu lassen, die Schallpegelkurve in annähernd logarithmischem Maßstab (dB), die lineare Schalldruckkurve und schließlich die eigentliche Tonhöhenaufzeichnung. Die Berandung zwischen hellem und dunklem Bereich gibt den Verlauf der Tonhöhe an. Am linken Rande ist der Frequenzmaßstab eingetragen. Die dicken horizontalen Linien entsprechen dabei den Oktaven des Tones c. Die Helligkeitsmodulation innerhalb des geschwärzten Aufzeichnungsbereiches rührt von den Obertönen (Formanten) her und ermöglicht die Abgrenzung von Vokalen und Konsonanten. Stimmlose Laute haben keinen Grundton und erscheinen daher wie Pausen als von oben nach unten durchgehende Linien.

Scharfe Aufzeichnungen ergeben sich indessen nur, wenn bei der Tonaufnahme genügende Sorgfalt aufgewandt wird. Insbesondere ist es notwendig, die Aufnahmen mit einem guten dynamischen oder Kondensator-Mikrophon in einem von Nebengeräuschen

\* *Grützmacher, M. und W. Lottermoser*: Akust. Z. 2, 1937, S. 242; *Kallenbach, W.*: Beihefte zur Acustica 1951, S. AB 37.

möglichst freien, schallgedämpften Raum vorzunehmen. Dabei ist auf eine gute Aussteuerung des Bandes zu achten, damit das Grundrauschen des Bandes möglichst ohne Wirkung bleibt. In den Abbildungen 5 bis 7 ist die Wirkung auf die Tonhöhenaufzeichnung dargestellt, die sich ergibt, wenn diese Maßnahmen nicht genügend beachtet werden.

Die häufigste Störungsquelle sind Brummspannungen durch Einstreuungen aus dem Leitungsnetz. Sie entstehen leicht bei längeren, hochohmigen Mikrophonleitungen, insbesondere wenn diese parallel zu Netzleitungen oder dicht am Netztransformator des Tonbandgerätes vorbei geführt werden. Auch eine ungenügende Siebung im Netzteil des Gerätes kann Brummaufzeichnungen zur Folge haben.

Die Wirkung auf die Tonhöhenaufzeichnungen ist in Abb. 5 zu erkennen. Hier wurde in einem Mischpult zu der Nutzaufzeichnung eine Brummspannung hinzugemischt, deren Größe jeweils unter den Teilbildern angegeben ist. Der Bezugswert 0 dB bedeutet dabei den Vollaussteuerungswert, der bei der Sprachaufzeichnung erreicht wird, -40 dB entspricht daher 1 %, -20 dB 10 % Brummszusatz. Es ist deutlich zu erkennen, daß die Brummspannung sich zunächst im komprimierten Oszillogramm bemerkbar macht, bei zunehmenden Amplituden ist sie auch in der Schallpegel- und Schalldruckkurve als Erhebung über dem Nullniveau erkennbar. Auf die Tonhöhenaufzeichnung wirkt sich die Brummstörung zunächst durch eine Girlandenstruktur der Berandungskurve aus, die die Frequenzbestimmung erschwert. Bei zunehmendem Brummpegel treten Ausreißer in der Aufzeichnung auf, die eine Auswertung schließlich unmöglich machen. Brummpegel von -5 dB werden in der Praxis zwar kaum eintreten, es soll hier jedoch deutlich zum Ausdruck gebracht werden, in welcher Form die Aufzeichnung darunter leidet. Werte von -15 bis -20 dB kommen dagegen gar nicht so selten vor. Bei diesen Pegeln ist die Qualität, besonders an schwächeren Stellen der Nutzaufzeichnung, schon erheblich gemindert.

Abb. 6 soll den Einfluß von Rauschanteilen zeigen. Solche Störungen treten ein, wenn der Aufnahmeraum nicht genügend gegen äußere Störquellen, z.B. Verkehrsgeräusche, abgeschirmt ist oder wenn das Tonband nicht genügend ausgesteuert wird, so daß das Grundrauschen des Bandes merklich in Erscheinung tritt. Zu bemerken ist auch, daß das Rauschen mit abnehmender Bandgeschwindigkeit ansteigt und sich beim Übergang von Vollspur auf



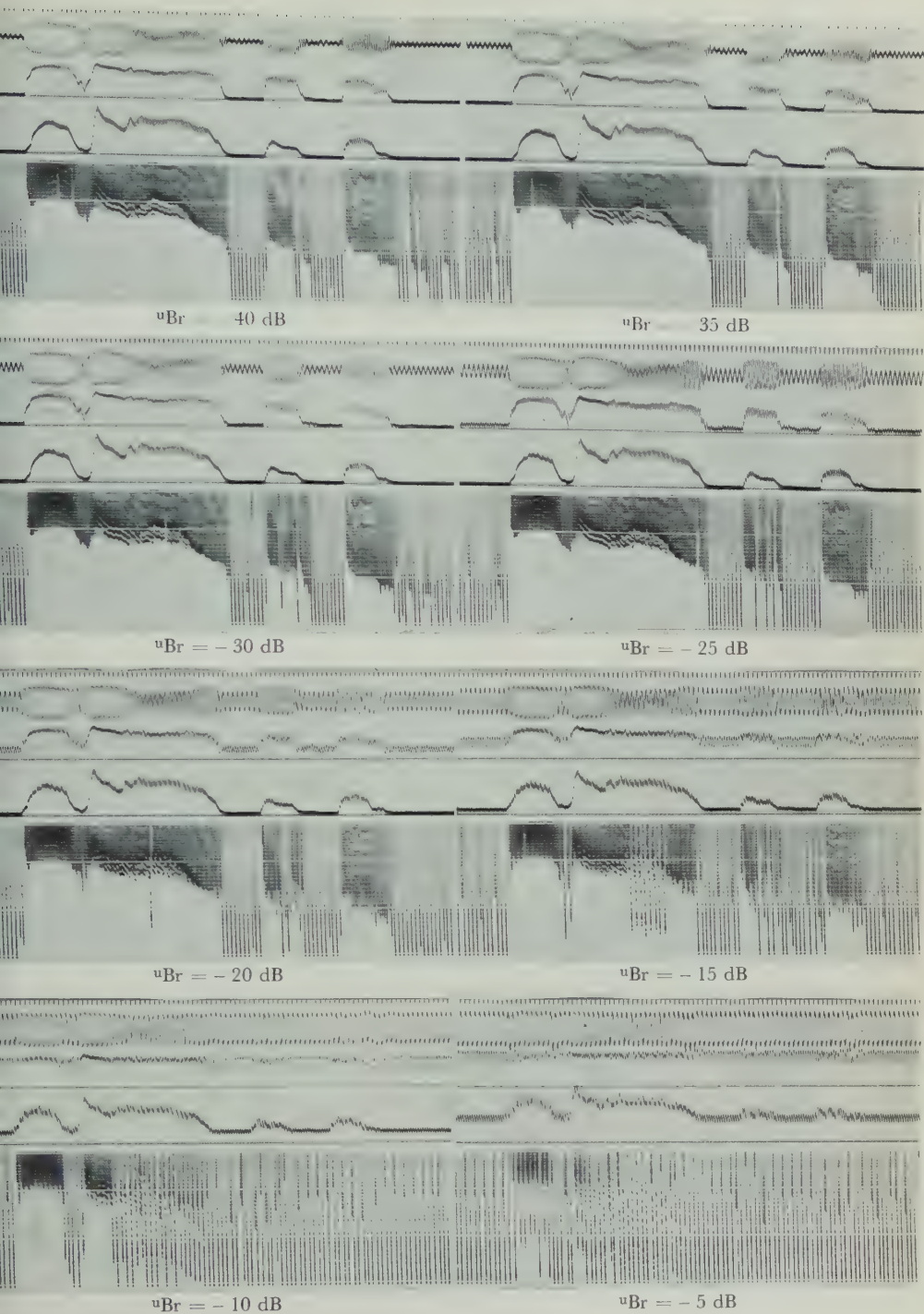


Abb. 5. Einfluß von Brummstörungen auf die Tonhöhenaufzeichnung.

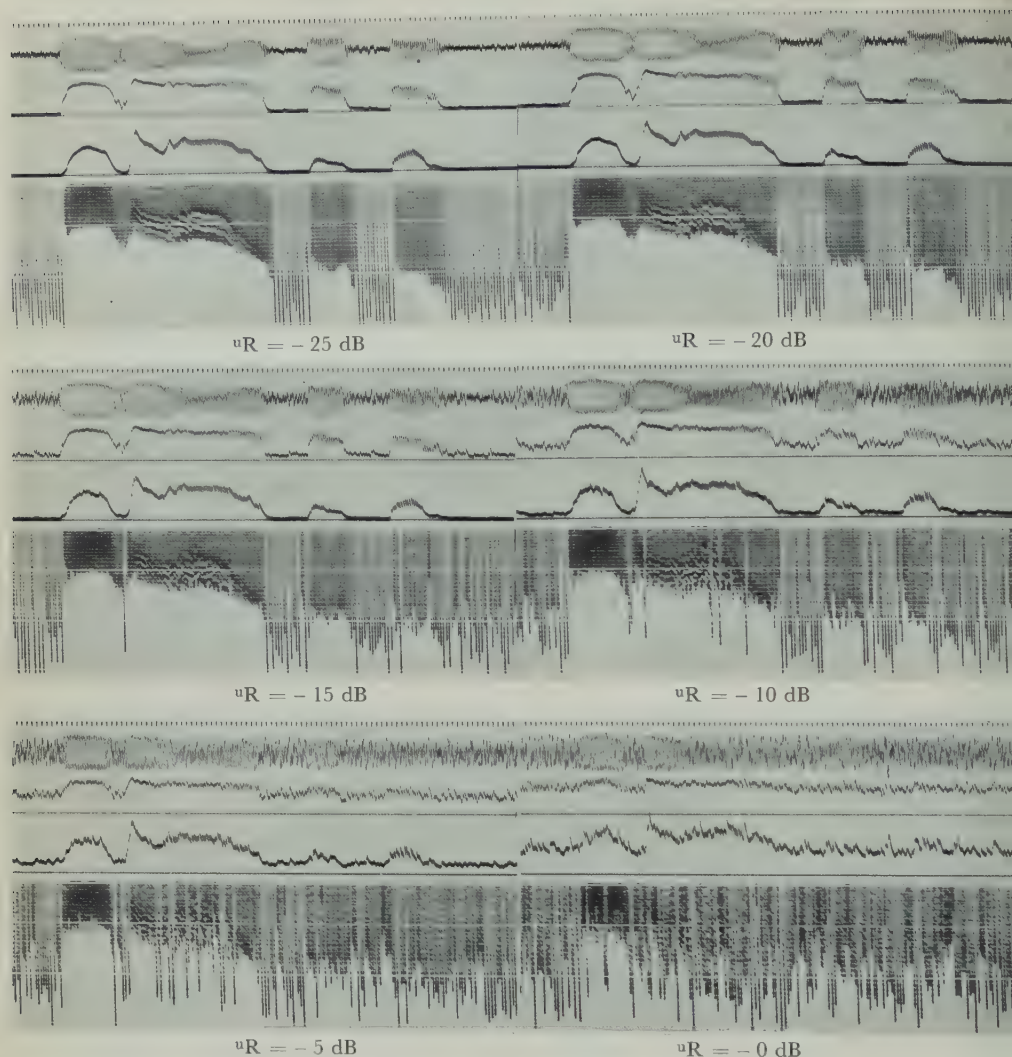


Abb. 6. Einfluß von Rauschstörungen auf die Tonhöhenaufzeichnung.

Halbspur um mindestens 3 dB erhöht, wie bereits oben erwähnt wurde.

Bei den Teilbildern von Abb.6 wurde der Nutzaufzeichnung eine Rauschspannung zugemischt, die einem Rauschgenerator entnommen wurde. Da bei natürlichen Rauschstörungen und auch im Bandrauschen die tieffrequenten Anteile gegenüber den hohen stärker enthalten sind, wurde die Rauschspannung noch mit einem

Filter bewertet, welches die tiefen Frequenzen stärker durchläßt als die hohen (« $\sqrt{\omega}$  = Gang»). Als Bezugswert 0 dB dient wiederum der Vollaussteuerungswert der Nutzaufzeichnung. Der Einfluß des Rauschens zeigt sich zunächst in der Unruhe der Nulllinie in den Sprechpausen, dann auch in den unregelmäßigen Schwankungen des Schallpegels und der Schalldruckkurve. Bei starkem Rauschen ist Nutz- und Störaufzeichnung nicht mehr zu trennen. In der Tonhöhenaufzeichnung ergibt sich mit zunehmendem Störgeräusch eine Ausfransung der Berandungskurve, die die Frequenzbestimmung erschwert und schließlich unmöglich macht. Am Schluß ist die Aufzeichnung völlig «verregnet».

Als letztes soll in Abb. 7 noch die Wirkung des Nachhalls und eines zu großen Mikrophonabstandes vom Sprecher dargestellt werden. Die Teilbilder wurden so gewonnen, daß die ursprüngliche Bandaufnahme, der die Abb. 4 entspricht, über einen Lautsprecher in einem etwa 20 m<sup>2</sup> großen möblierten Laboratoriumsraum abgespielt wurde, der weder Gardinen noch sonstige schallschluckende Verkleidungen enthielt. Vor dem Lautsprecher wurde nun das Mikrophon in verschiedenen Abständen  $a$  aufgestellt und das Klangbild erneut auf Tonband aufgenommen. Aus diesen Aufnahmen wurden die Teilbilder von Abb. 7 gewonnen.

Es muß betont werden, daß auch bei kürzestem Mikrophonabstand (0,25 m) schon mit gewissen Verzerrungen durch den Frequenzgang des Lautsprechers zu rechnen ist. Dennoch ist bei dem Oszillogramm, der Pegel- und Schalldruckkurve zunächst noch eine weitgehende Ähnlichkeit mit Abb. 4 zu erkennen. Mit zunehmendem Mikrophonabstand verändern sich die Kurven unter dem Einfluß des Nachhalls und der Raumresonanzen immer stärker, so daß schon bei 2 m Abstand erhebliche Verzerrungen eintreten. Ferner ist zu erkennen, daß die Laute vor den Plosivlauten nicht mehr wie in Abb. 4 ziemlich plötzlich abbrechen, sondern infolge des Nachhalls einen Schlauch hinter sich herziehen, der sich exponentiell verengt. In der Tonhöhenaufzeichnung zeigt sich schon bei kurzem Abstand eine deutliche Veränderung des Modulationscharakters durch den Nachhall, wobei allerdings auch die Frequenzkurve des Lautsprechers von Einfluß ist. Mit zunehmendem Abstand franst auch hier die Berandungskurve aus, so daß die Auswertung erschwert wird. Die Zerstörung der Information ist jedoch nicht so stark wie bei den Brumm- und Rauschstörungen.

Die Dämpfung des Aufnahmeraumes durch Gardinen, Teppiche



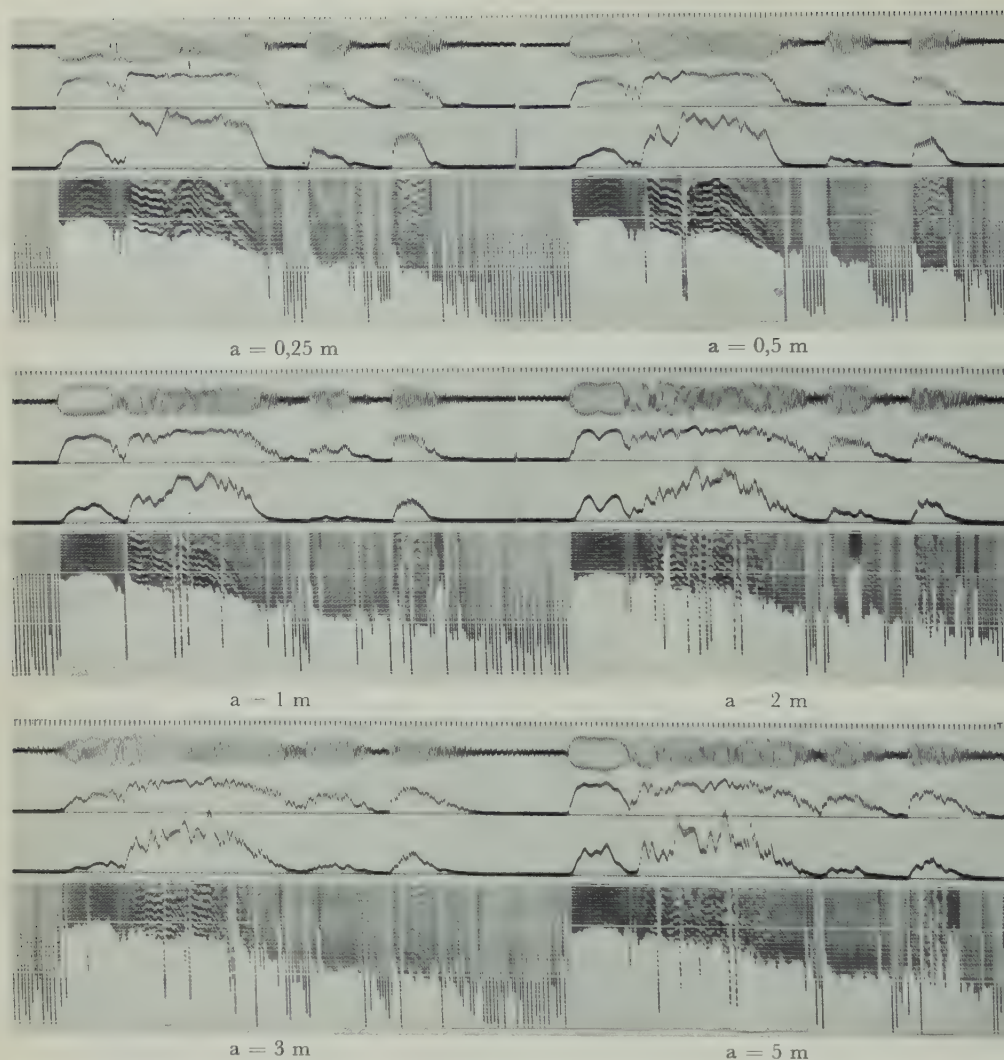


Abb. 7. Einfluß des Mikrophonabstandes  $a$  von der Schallquelle auf die Tonhöhenaufzeichnung.

oder Polstermöbel ist vor allem dann von Bedeutung, wenn aus besonderen Gründen, etwa um die Versuchsperson nicht zu beeinflussen, das Mikrophon versteckt in größerem Abstand aufgestellt werden muß.

Zum Schluß kann zusammenfassend gesagt werden, daß bei dem heutigen Stande der Technik hohen Ansprüchen genügende

Aufnahmen bei 19 cm/s (Halb- oder Vollspur) ausgeführt werden können. Auch mit 9,5 cm/s Bandgeschwindigkeit lassen sich noch in vielen Fällen brauchbare Aufnahmen erzielen, wenngleich die oben geschilderten Gefahren beim Bandaustausch schon beträchtlich ansteigen. Bei der Auswahl der Geräte sollte man seine Aufmerksamkeit besonders auf die «semiprofessionellen» richten, die vielfach nahezu Studioqualität besitzen und im Preise nicht allzu weit über den Heimgeräten liegen. Ebenso wichtig wie das Aufnahmegerät selbst ist jedoch der Aufnahmeraum und die Beachtung der geschilderten Maßnahmen bei der Aufnahme.

### *Zusammenfassung*

Die Frage, welches Tonbandgerät und welche Bandgeschwindigkeit bei Sprachuntersuchungen zu wählen sind, hängt von der gestellten Aufgabe ab. Es wird gezeigt, worin die Qualitätsminderung bei Verwendung von Heimgeräten mit kleiner Bandgeschwindigkeit besteht. Sie liegt in der Zunahme der Pegelschwankungen, besonders bei hohen Frequenzen sowie in der Schwierigkeit, beim Austausch von Bändern den Spaltschiefstellungseffekt zu vermeiden. Sollen von Bandaufnahmen Tonhöhenaufzeichnungen angefertigt werden, so ist bei der Aufnahme auf Brummfreiheit, Vermeidung von Störgeräuschen und Nachhall und genügende Bandaussteuerung zu achten. Die Wirkung dieser Einflüsse wird anhand von Probeaufzeichnungen dargestellt. Für hochwertige Aufnahmen wird die Bandgeschwindigkeit 19 cm/s und die Verwendung «semiprofessioneller» Geräte empfohlen.

### *Tape Recording Techniques in Voice Research*

#### *Summary*

The question of which tape recorder and which tape speed to choose for voice research depends on the intended requirement in this field. It is shown that a reduction of quality results from the use of "home" type equipment having a low tape speed. This reduction of quality is due to large level variations, especially in the high-frequency range and the difficulty in avoiding the effect of misalignment of the gap when changing tapes. If it is required to extract precise frequency levels from the tape, it is necessary to provide hum suppression, eliminate background noises, and reverberation, and maintain an adequate modulation level. The effect of these factors is represented on test recordings. For high-quality recordings a tape speed of 19 cm/sec and the use of semi-professional apparatus is recommended.

*Remarques sur la technique d'enregistrement d'études phonétiques**Résumé*

Le problème du choix d'un type de magnétophone et d'une vitesse d'enregistrement se pose, lors d'études phonétiques, en fonction du sujet. Il est montré qu'une diminution de qualité résulte de l'utilisation de magnétophones dits d'amateur, dont la vitesse d'enregistrement est faible. Cette diminution de qualité est conditionnée par de plus grandes variations du niveau sonore, notamment dans le domaine des hautes fréquences, ainsi que par la difficulté d'éviter l'«effet de décalage à l'interstice» lors du changement de bande. S'il s'agit de procéder à des analyses de hauteur du son, il faut veiller à ce que les ronflements et bruits parasites, ainsi que les phénomènes de réverbération soient éliminés. En outre, il faut maintenir un niveau d'enregistrement constant. Les effets de ces facteurs sont représentés à l'aide d'enregistrements d'essai. Si l'on veut effectuer des enregistrements de haute qualité, il conviendra d'utiliser des appareils «semi-professionnels» et la vitesse de 19 cm/sec.

Adresse der Autoren: Dr. Werner Kallenbach und Dipl.-Ing. Hans-Jürgen Schroeder, Physikalisch-Technische Bundesanstalt, Bundesallee 100, Braunschweig (Deutschland)



## Sur le problème de la syllabe

### Réponse à M. B. Hála

Par A. ROSETTI, Bucarest

M. Hála, dans un article publié ici-même<sup>3</sup>, se propose de combattre notre manière d'envisager la nature de la syllabe.

Nous allons examiner successivement les arguments de M. Hála, qui, à notre avis, ne sont pas de nature à ruiner notre explication.

Mais il convient, tout d'abord, de rappeler que notre argumentation est en partie fondée sur les recherches de phonéticiens qui font autorité, tels que *R. H. Stetson*, *Ch. F. Hockett*, *E. Haugen*, *N. I. Žinkin* et d'autres, ou de linguistes tels que *K. Bühler*, chose que M. Hála passe sous silence. Nos résultats coïncident avec les résultats des recherches des savants cités ci-dessus. M. Hála affirme avec force que notre formule «il n'y a pas de syllabe sans air» est fausse, mais il oublie de réfuter les résultats des recherches de tous les autres phonéticiens, qui sont fondées, cependant, sur une masse de faits.

Qu'y-a-t-il de fautif, dans les expériences, analyses et jugements cités dans notre mémoire sur la syllabe (paru en 1959) et qui fait l'objet des remarques de M. Hála, voilà, sans doute, ce que M. Hála aurait dû nous montrer tout d'abord.

Au lieu de la réfutation attendue, M. Hála se livre à une série de remarques qui évitent le sujet essentiel de la controverse et que nous allons examiner maintenant.

1° M. Hála s'élève contre notre classification de «phonèmes à air» et «phonèmes sans air». Pour nous convaincre, il prend comme exemple la consonne *p* qui, à son avis, fait partie des phonèmes à air, puisque il y a une «petite portion d'air s'échappant à la détente» de cette consonne (p. 160).

Cette remarque doit-elle être prise au sérieux?

Car qui pourrait nier que les consonnes occlusives sont des phonèmes fermés, en regard des voyelles ou des consonnes non-occlusives?

2<sup>o</sup> M. Hála critique notre classification des phonèmes en «phonèmes à air», ou ouverts, opposés aux «phonèmes sans air», ou fermés. Il cite à ce propos une remarque critique de M. F. Hintze, à laquelle nous avons répondu ici-même<sup>10</sup>, en montrant que les semi-voyelles ou les semi-consonnes qui, du point de vue de leur fonction, jouent dans la syllabe le rôle de consonnes, sont *en même temps* des phonèmes ouverts, du point de vue acoustico-articulatoire, et des phonèmes fermés, du point de vue fonctionnel.

3<sup>o</sup> Si les phonèmes *m, n, l, r, s, ʃ, f* peuvent former syllabe, pourquoi *x, h, v, z, ʒ, j, w* ne forment pas syllabe, objecte M. Hála (p. 162).

Cette objection n'est cependant pas de nature, à renverser notre théorie. Tout d'abord parce que la langue n'emploie pas *tous* les moyens dont elle dispose, et il y a maintes combinaisons possibles de phonèmes qui sont restées sans emploi\*. Et ensuite parce qu'il y a des langues non-indo-européennes qui forment des syllabes avec ces éléments. Ainsi, en Bella Coola (Amérique du nord), langue sans voyelles, il y a des mots formés par des syllabes de la structure suivante: *l-k'w-t-xw* «make it big!», *k'-x-t-c* «I looked» ou bien *s-k'l-xk-x-c* «I'm getting cold»\*\*.

M. Hála nie la valeur démonstrative des syllabes telles que fr. *pst, cht*, etc.\*\*\*. «En réalité», affirme M. Hála, «il ne s'agit pas ici de syllabes, mais tout simplement de bruits servant de signaux expressifs et que l'on cherche à traduire, dans la langue écrite, par des lettres correspondant plus ou moins à l'effet acoustique des dits signaux. Mais ces derniers, connus en grammaire sous le nom d'*interjections* ou d'*onomatopées*, relèvent d'autres lois que les mots normaux» (p. 162).

Nous ne partageons pas ce point de vue.

\* Voir Rosetti<sup>9</sup>, p. 26.

\*\* Voir Hockett<sup>4</sup>, p. 57.

\*\*\* M. Hála (p. 162) se demande «pourquoi l's devrait-il former syllabe p. ex. dans le tchèque *pst* et non pas dans le tchèque *sto*?». Parce que *s* n'est pas *toujours* syllabique. C'est ce qui arrive, par exemple, pour *l* tchèque, qui est syllabique dans *vlk* «loup» et non-syllabique dans *vliv* «influence».

Tout d'abord, parce que les onomatopées, comme l'a montré *Maurice Grammont*, imitent d'une manière *approximative* les sons naturels, qui n'appartiennent pas à la langue parlée, parce qu'ils font partie d'une série différente\*.

Les onomatopées *sont donc faites avec des sons parlés*. C'est pourquoi les onomatopées rentrent, pour la plupart, dans le vocabulaire de la langue donnée et qu'elles deviennent des mots de la dite langue. (Le fait que les dictionnaires enregistrent ces formations parmi les mots du vocabulaire de chaque langue, confirme cet argument.)

Il est facile d'apporter des preuves à l'appui de cette affirmation, qui renverse la thèse de *M. Hála*.

Prenons, par exemple, les «Lautbilder» des langues négro-africaines\*\*, qui forment des mots de la langue courante. Ainsi *hèhè-hèhè*, en ewe, «beschreibt einen Korpulenten, etwas schwerfällig und schwabbelig einhergehenden Menschen», *gblugblugblu* «ist der Gang des stier und steif vor sich hingehenden Büffels oder eines entsprechend gebauten Menschen», etc. *D. Westermann* a réuni un grand nombre de mots ainsi formés (op. cit., p. 166 et s., 194 et s.).

Le roumain offre de nombreux exemples similaires. Des verbes tels que roum. *a bîjbîi* «chercher à tâtons», *a mormăi* «murmurer», etc. sont d'origine onomatopéique: ici aussi, la langue a incorporé les onomatopées dans son vocabulaire courant\*\*\*.

4° *M. Hála* s'efforce d'éliminer notre argument fondé sur la voix chuchotée, émission phonatoire qui connaît la division syllabique, malgré l'absence de la sonorité laryngienne.

*M. Hála* s'accroche, tout d'abord, à une question de terminologie: il n'y a pas de «voix» sans «son» nout dit-il (p. 165), et personne ne le contredira là-dessus. Disons, cependant, que le terme de «voix chuchotée» est d'un emploi courant. Voici quelques citations à l'appui de notre affirmation: «la *voix chuchotée* diffère de

\* Voir *Grammont*<sup>1</sup>, p. 97 et s.; *Grammont*<sup>2</sup>, p. 377 et s.

\*\* *Westermann*<sup>13</sup>, «Ein Lautbild», nous dit l'auteur, «ist ein Wort, das im Empfinden der Sprachgemeinschaft eine lautliche Reaktion auf einen empfungenen Sinnesindruck darstellt, ein Wort also, das einem inneren Gefühl einen unmittelbaren, lautlich wie sprachmelodisch angemessenen Ausdruck gibt» (p. 159).

\*\*\* Voir *Pușcariu*<sup>7</sup>, p. 105. On consultera avec profit l'exposé de l'auteur, *Les onomatopées de la langue roumaine*, dans *Etudes de linguistique roumaine*<sup>3</sup>. Le grand nombre de mots, dérivés d'onomatopées, en roumain, est illustré par l'exemple suivant: dérivés de *bîză* «mouche, abeille»: *bîzii* (*bîziire*, *bîziit*, *bîzîtură*, *bîziială*, *bîziitor*), *bîzîitoare*, *bîzălău*, *bîzalnic*, *-ă*, *bîzău* (p. 349-350).



la voix haute»\*; «si la partie de la glotte qui se trouve entre les aryténoïdes reste ouverte en laissant passer l'air, on obtient une *voix chuchotée*\*\* ; *chuche* ou *voix chuchée*: émission de voix... La *voix chuchée* se rapproche ainsi plutôt de la voix murmurée, que de la *voix chuchotée*\*\*\*.

Mais peut-on nier l'existence réelle de la syllabe dans cette émission non-sonore? Comme la chose est indéniable, M. *Hála* s'en tire en nous disant que le chuchotement est un langage «transposé à un autre plan de la réalisation communicative» (p. 166), ce qui nous plonge dans le vague (quel plan?!) et laisse le problème ouvert.

Nous considérons, par conséquent, que notre argument est pleinement valable: si la sonorité laryngienne, modifiée dans les cavités sus-glottiques et buccale n'est pas essentielle pour la formation de la syllabe, il reste que les efforts de M. *Hála* pour ruiner la théorie de la syllabe fondée sur la présence de l'air phonateur, sont restés vains.

Que dire, enfin, de la «définition d'ordre acoustique» de la syllabe, proposée par M. *Hála*, que nous reproduisons ici: «la syllabe serait l'élément fondamental de la parole produit par chaque mise en phonation (résonnement, retentissement) isolée de la voix laryngienne modifiée au passage des cavités sus-glottiques et notamment de la cavité buccale sous forme de diverses sonantes (voyelles, diphtongues, consonnes appelées par moi «sonores»: liquides et nasales»; p. 164), si ce n'est que cette définition est confuse et qu'elle est formulée en des termes impropres. Car qu'est-ce qu'une émission *isolée* (quand on sait que les sons du langage, dans la chaîne parlée, se suivent d'une manière ininterrompue); et qu'est-ce, encore, que la «voix laryngienne... modifiée... *sous forme* de diverses sonantes», etc.?

Convenons-en, la tentative de M. *Hála* de donner une définition de la syllabe acoustique a échoué.

### Résumé

En réponse à l'article dans lequel M. B. *Hála* s'est proposé de ruiner notre théorie de la syllabe, fondée sur la présence de l'air, M. *Hála* donne une définition acoustique confuse de la syllabe, que l'on

\* Voir Rousselot<sup>11</sup>, p. 468: «Voix chuchotée».

\*\* Voir Malmberg<sup>6</sup>, p. 22, et aussi fig. 20, C, p. 22.

\*\*\* Voir Marouzeau<sup>6</sup>, p. 53.

ne saurait retenir. M. *Hála* omet de réfuter les expériences et arguments des phonéticiens qui ont étudié le côté physiologique du problème. Il est montré, ensuite, que la tentative de M. *Hála* de réfuter nos arguments concernant les phonèmes ouverts et fermés, les onomatopées et le chuchotement a échoué.

### *Zum Problem der Silbe*

#### *Zusammenfassung*

Bei dem Versuch von B. *Hála*, die Silbentheorie des Verfassers, die auf der Beteiligung der Luft beruht, zu Fall zu bringen, wird eine verworrene akustische Definition der Silbe vorgelegt, die nicht aufrechtzuerhalten ist. *Hála* unterläßt es, die Experimente und Argumente der Phonetiker zu widerlegen, die die physiologische Seite des Problems untersuchten. Es wird dann gezeigt, daß der Versuch *Hálas* gescheitert ist, die Argumente des Verfassers im Hinblick auf die offenen und geschlossenen Phoneme, die Onomatopoetica und das Flüstern zu entkräften.

### *On the Problem of the Syllable*

#### *Summary*

Intending to fell the author's theory of syllables based on the presence of air, B. *Hála* presented a confused acoustical definition of the syllable, which is not tenable. *Hála* omitted to disprove the experiments and arguments of phoneticians having studied the physiological view of the problem. It is shown that *Hála* failed in refuting the author's arguments concerning open and closed phonemes, onomatopoetica, and whisper.

### *Bibliographie*

1. Grammont, M.: Onomatopées et mots expressifs. Rev. langues romanes 46: 97-158 (1901).
2. Grammont, M.: Traité de phonétique (Delagrave, Paris 1933).
3. Hála, B.: Autour du problème de la syllabe. *Phonetica* 5: 159-168 (1960).
4. Hockett, C. F.: A manual of phonology (Waverly Press, Baltimore 1955).
5. Malmberg, B.: La phonétique (Presses universitaires de France, Paris 1954).
6. Marouzeau, J.: Lexique de la terminologie linguistique. 2<sup>e</sup> éd. (Geuthner, Paris 1951).
7. Puşcariu, S.: Die rumänische Sprache (Otto Harrassowitz, Leipzig 1943).
8. Puşcariu, S.: Etudes de linguistique roumaine (Cluj-Bucureşti 1937).
9. Rosetti, A.: Le mot (E. Munksgaard, Copenhagen-Bucarest 1947).
10. Rosetti, A.: Sur la classification des phonèmes semi-voyelles (ou semi-consonnes). *Phonetica* 5: 123-124 (1960).
11. Rousselot, P. J.: Principes de phonétique expérimentale. 2<sup>e</sup> éd. (Didier, Paris 1924).
12. Westermann, D.: Laut und Bild in einigen westafrikanischen Sprachen. Arch. vergl. Phon. 1: 154-172, 193-212 (1937).

Adresse de l'auteur: Professeur A. Rosetti, 56, Str. Dionisie Lupu, Bucarest 36 (Roumanie)

From the Communication Sciences Laboratory, University of Michigan, Ann Arbor, Mich.

## Some Acoustic Correlates of Accent in Serbo-Croatian

By ILSE LEHISTE

### 1. *Introduction*

The Serbo-Croatian standard language is characterized by the presence of features from three partially overlapping prosodic systems. Contrasts between long and short vowels may occur in both accented and unaccented position; quantity thus appears as an independent prosodic system in the standard language. Both stress and pitch are involved in the accentual patterns, which are usually described as rising and falling. A combination of these with the two vowel quantities yields a system comprising four accent types: short falling, long falling, short rising, and long rising. There are important restrictions in the distribution of the four accent types. All monosyllabic words have falling accents; in polysyllabic words, falling accents are restricted to occurrence on the first syllable, whereas rising accents may occur on any syllable except the last.

In the dialect spoken by the educated population of Belgrade, a change in the distribution of these accents appears to have taken place during the past two or three decades. *Meillet and Vaillant*<sup>1</sup> note that the two short accents tend to merge, or rather that the short rising accent appears to be assimilated to the short falling accent. *Belić*<sup>2</sup> states that the short rising tone in two-syllable words with a short second syllable is often replaced by the short falling tone. The contrast is preserved in words with a long vowel in the second syllable. Another indication of the instability of the short rising accent is the fact that a long rising accent is often substituted for the short rising accent.

### 2. *Previous phonetic descriptions*

A survey of both traditional and experimental phonetic studies



of the Serbo-Croatian accents is given by *Pollok*<sup>3</sup>. According to publication date, the most recent of these is a paper by *Appel*<sup>4</sup>. However, the work reported by *Appel* was done in 1927, and instrumental methods have been greatly refined since then. More recent instrumental work on Serbo-Croatian phonetics is described in a paper by *Kostić*<sup>5</sup>. This paper emphasizes intensity measurements, and deals only with the two short accents. It appears, therefore, that there exists no up-to-date experimental phonetic description of both the fundamental frequency contours and intensity patterns of the four accent types of standard Serbo-Croatian.

The phonetic realization of the accents is usually briefly touched upon in papers dealing with the phonology of Serbo-Croatian. In phonemically oriented papers, there appears to be a tendency to evaluate the phonetic observations by assigning primacy to either the intensity (stress) or fundamental frequency (pitch) features. The description offered by *Lüdtke*<sup>6</sup> may be considered representative of the first point of view. The description given by *Rešetar*<sup>7</sup> is based purely on pitch features. *Meillet and Vaillant*<sup>8</sup> state explicitly that pitch features rather than intensity have to be considered significant. In the phonetic description contained in *Belić's* work<sup>9</sup>, both intensity and pitch features are considered, neither being given preference over the other. Agreement does not seem to exist either about the phonetic nature of the accentual phenomena involved, nor about their possible phonemic interpretation. Thus there appears to be some justification for a new attempt to establish the phonetic characteristics of the four accent types by the relatively more objective techniques of experimental phonetics.

### *3. The present investigation*

The study reported here was undertaken with the purpose of providing an acoustic-phonetic description of the accent types used by speakers of the dialect of Belgrade.

#### *3.1. Informants, material, and method*

The speech of three informants<sup>10</sup> was analyzed for the study. All three informants belong to the younger generation and have received their education – from elementary school through the university – in Belgrade. Of the informants, DmK recorded a set of words incorporating minimal tonal contrasts, short sentences con-

taining the words included in the contrastive pairs, and a longer connected text. DK recorded a set of 100 frame sentences. The 100 words used in the frame contained 29 words with short falling accent, 41 words with long falling accent, 13 words with short rising accent, and 17 words with long rising accent. VJ recorded a set of sentences, examples of poetry, and a short continuous text. In addition, DK listened to the tapes recorded by DmK, and VJ listened to the tapes of the two other informants, identifying the words and offering further comments. The data thus consist of both actual recordings and listener judgments about the test material.

The recordings were made in the Communication Sciences Laboratory of The University of Michigan, using a high quality tape recording system. The tapes were analyzed by spectrographic techniques, using the Model D sound spectrograph available at the Laboratory. Four inch narrow-band spectrograms (with a 45 cps filter) and two inch broad-band spectrograms (with a 300 cps filter) were made of the test utterances. The tapes were further processed through a speech power measuring circuit developed at the Laboratory<sup>11</sup>. Measurements made from these instrumental analyses constitute the basis from which the patterns were derived that are described in this paper. Since the material recorded by DK lends itself best to quantitative treatment, the main part of the paper will report measurements made from the speech of this informant.

The list of words recorded by DK was compiled to include at least one occurrence of each vowel and tone combination. The 100 words were pronounced in the frame "Fráza . . . je nàpīsānā na tábli". The frame was selected such that words belonging to different categories and occurring in different inflectional forms might be commuted in the same slot. When the recording was made, a randomized list was given to the informant, containing the Serbian words in simple transliteration, and an English translation indicating the desired form. Informant DK selected the Serbian form that corresponded, in his usage, to the English word. The list of words is given in the Appendix.

The hundred repetitions of the frame sentence offered an opportunity both to analyze 100 different words in the commutation slot, and to study 100 occurrences of the words constituting the frame. The 100 occurrences of the word *fráza* were also analyzed; it is thus possible to make some generalizations about the range of variations possible within a number of repetitions of the same pattern.

Table I

Fundamental Frequency Patterns in Test Words (Frequency in cycles per second)

Test word type	First syllable					Second syllable			Third syllable		
	No.	Beg.	Peak	End	Peak occurs at % of total duration	No.	Beg.	End	No.	Beg.	End
Short falling (29)											
Monosyllabic	11	122	127	109	46%						
	(-3r)										
Disyllabic	14	115	127	122	70%	15	114	93			
	(-1r)					(-5 lar.)					
Trisyllabic	4	115	133	129	70%	4	116	105	4	95	92
Long falling (41)											
Monosyllabic	21	122	130	89	24%						
	(-5r)										
Disyllabic	20	126	139	94	30%	20	87				
	(-1r)					(-1 lar.)					
Short rising (13)											
Disyllabic	8	115	127	124	86%	8	121	96			
	(-1r)					(-1 lar.)					
Trisyllabic	5	105	122	120	83%	5	129	116	5	97	
	(-4r)								(-1 lar.)		
Long rising (17)											
Disyllabic	16	107	122	121	97%	16	121	96			
	(-1r)										
Trisyllabic	1r					1	130	110	1	110	100
100 occurrences of <i>fráza</i>	100	104	123	121	91%	100	116	101			

### 3.2. Fundamental frequency patterns

Table I contains information about the fundamental frequency patterns observed on the 100 words used in the frame, and the 100 occurrences of the first word in the frame. The following measurements are included in the table. The fundamental frequency on the first syllable of the test word was measured at the beginning of the vowel, at the peak of the fundamental frequency contour, and at the end of the vowel. In words of more than one syllable, the fundamental frequency was also measured on the successive syllables. The moment in time at which the peak occurred was ascertained and expressed in terms of per cent of the total duration of the vowel. Differences in the duration of the syllable nuclei could be compensated for in this manner. The values in the table refer to averages computed for the number of words in each classification that were



contained in the set of analyzed words. The second column indicates the number of words included in each row. The fundamental frequency on words with /r/ as syllable nucleus was not included in the averages for the first syllable, but the second syllables of such words were included in computing the averages for those syllables. The syllable nucleus /r/ presents special problems in measurement; this syllable nucleus will be treated separately in a later section. On the other hand, the fundamental frequency was not measured on those unstressed syllables where the speaker used laryngealized voice quality (i.e. where the pitch dropped so low that the mode of phonation changed and the vibrations became irregular).

The patterns occurring on long vowels will be considered first. The long falling pattern was observed on 21 monosyllabic words and 20 disyllabic words. In these test words, the pattern on the monosyllabic words was practically identical with the pattern on the first syllable of disyllabic words. On the average, the fundamental frequency reached a peak during the first third of the accented vowel in both cases; after attaining peak value, the fundamental frequency was falling throughout the syllable. An inspection of the spectrograms indicates that the beginning of the fundamental frequency movement on the test words was conditioned by the phonetic character of the preceding consonant: the fundamental frequency started high after a voiceless sound, but rose from a lower frequency value after a voiced sound<sup>12</sup>. Thus the beginning value of the contour is irrelevant. The total fall in frequency occurred on the syllable bearing the accent; the second syllable of disyllabic words had a steady fundamental frequency at approximately the low end of the speaker's voice range, and there was one instance of laryngealization. The ratio of the frequencies between peak and terminal frequency corresponded approximately to a musical fifth.

The rising pattern on long vowels was observed on 16 disyllabic words, 1 trisyllabic word, and on 100 occurrences of the word *fráza* as the first word of the frame sentence. The words showed a consistently rising pattern of fundamental frequency, starting at a frequency considerably lower than the starting frequency in words with the long falling accent, and reaching their peak immediately before the consonant. On the average, the rise involved an interval smaller than a minor third. The second syllable started at a frequency close to the terminal frequency of the first syllable, and then dropped approximately a major third. The drop in the second syllable of

*fráza* was approximately equal to a minor third. The difference by one semitone is attributable to the fact that the first word in the sentence was subjected to a rising intonation which reached its peak in the test word. The single trisyllabic word in the set had a fundamental frequency pattern on its second syllable that differed from the second syllable patterns of disyllabic words, and resembled one that occurred on five trisyllabic words with short rising tone. On these words, the second syllable started at a frequency noticeably higher than the terminal frequency of the first syllable, and the fall was distributed over the second and the third syllables. This pattern will be described in more detail in connection with the short rising patterns.

Figure 1 contains reproductions of narrow-band spectrograms of six utterances by informant DK, illustrating long rising and long falling patterns. On the left hand side are spectrograms of three words with long falling accent, preceded by the word *fráza*. On the right hand side are three words with long rising accent. The tenth harmonic has been traced with white paint on the spectrograms, and provides a visual representation of the fundamental frequency movement. In cases where the tenth harmonic became too weak to be followed, the fifth harmonic has been traced with a dotted line. It is apparent that the fundamental frequency started high in the word *súd*, where the stressed vowel was preceded by a voiceless fricative, but that it started at a lower frequency and reached a peak near the beginning of the vowel in the words *bôj* and *Lúka*. The segment corresponding to /j/ in *bôj* and the vowel of the second syllable in *Lúka* were laryngealized, and the 45 cps filter of the spectrograph was not able to resolve the sound wave into separate harmonic components. The words with long rising accent show the characteristic distribution of the fundamental frequency pattern over the two syllables.

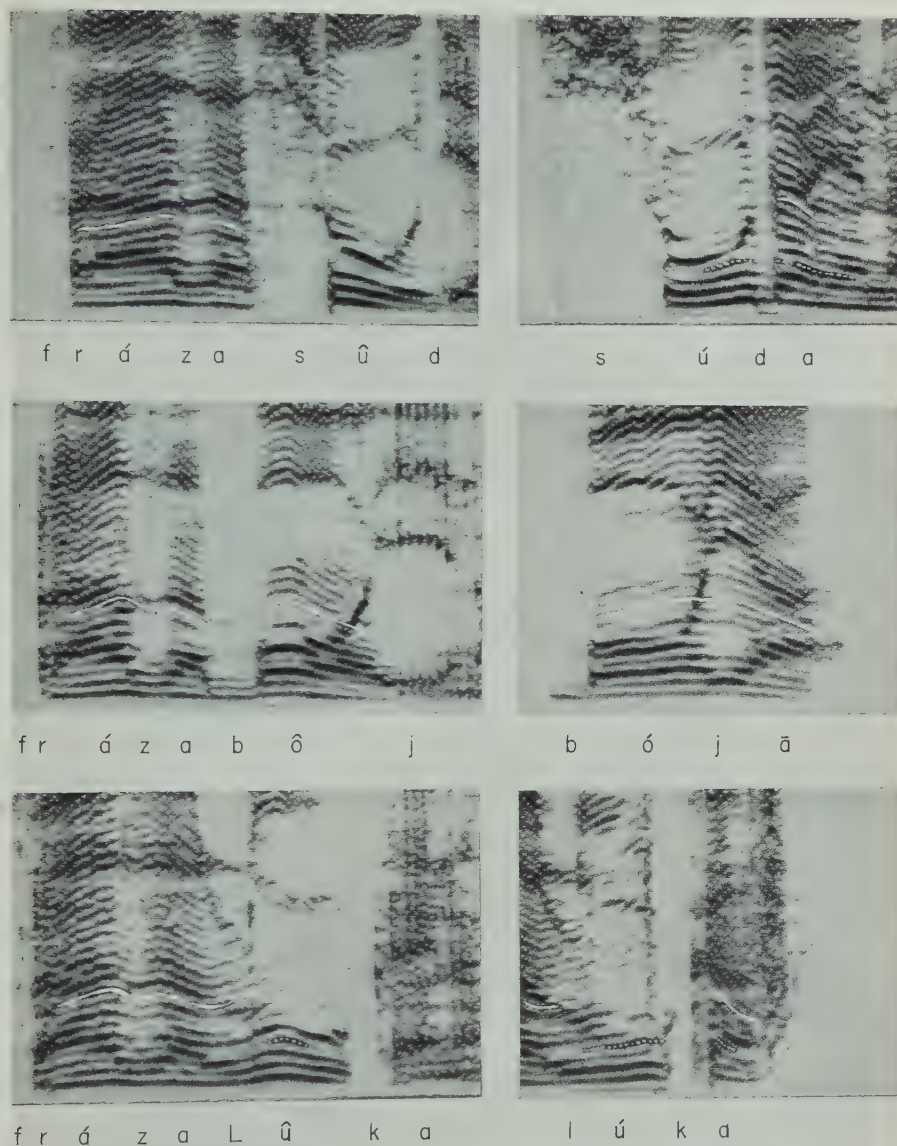
The falling pattern on short vowels was observed in 11 monosyllabic words, 14 disyllabic words, and 4 trisyllabic words. The short falling pattern differed from the long falling pattern in the distribution of the fundamental frequency contour over the test word. In monosyllabic words with the short falling pattern, the peak in the fundamental frequency was reached approximately in the middle of the test word, and the drop in frequency was somewhat less than a minor third. In disyllabic words, the peak was reached toward the end of the stressed vowel (approximately  $\frac{2}{3}$  to  $\frac{3}{4}$  of the

duration), and the fall took place on the second syllable, although the frequency at the beginning of the second syllable was lower than in words with long rising accent, which had a similar fall on the second syllable. In trisyllabic words, the drop in frequency was distributed over the second and third syllables. There were five instances of laryngealization in the second syllable of disyllabic words.

The words in the next category, the short rising accent, were selected to include a number of instances where according to the handbooks a short rising accent should occur<sup>13</sup>. The informant was not instructed to produce a word with short rising accent, but rather to give the Serbian equivalent of an English word given in the form in which the standard Serbo-Croatian equivalent should exhibit a short rising accent. An analysis of these words, as well as further observations made with connected speech of all informants, suggests that there is very little reason to treat the two-syllable words in this group as different from the words included in the short falling category. The fundamental frequency pattern occurring on 8 disyllabic words included in this group was very similar to that occurring on the disyllabic words with short falling accent. On the average, the fundamental frequency reached its peak slightly later in the first syllable of disyllabic words with short rising accent (at 86% of the total duration of the stressed syllable nucleus) than in disyllabic words classified as belonging to the short falling category (at 70% of the total duration). The second syllable of the words with short rising accent began at a slightly higher fundamental frequency than the second syllable of the words with short falling accent; however, there was also one instance of laryngealization in the second syllable of words classified as having a short rising accent.

The 5 trisyllabic words where a short rising accent was expected differed more markedly from corresponding words with a short falling accent. The fundamental frequency on the second syllable of the trisyllabic words with short rising accent started higher than the peak in the first syllable, and dropped only slightly during the second syllable. The third syllable was produced at a low frequency, and there was one instance of laryngealization in the third syllable. The conclusion to be drawn from these limited data is that no difference exists in the fundamental frequency patterns of disyllabic words with short rising and short falling accent, but that there is a difference associated with words of three syllables.





*Fig. 1.* Narrow-band spectrograms of six utterances by informant DK. The tenth harmonic has been traced with white paint; on some of the spectrograms, the fifth harmonic has been indicated with a dotted line.

Figure 2 illustrates the fundamental frequency patterns associated with words classified as having short rising and short falling accents. Narrow-band spectrograms, broad-band spectrograms, and

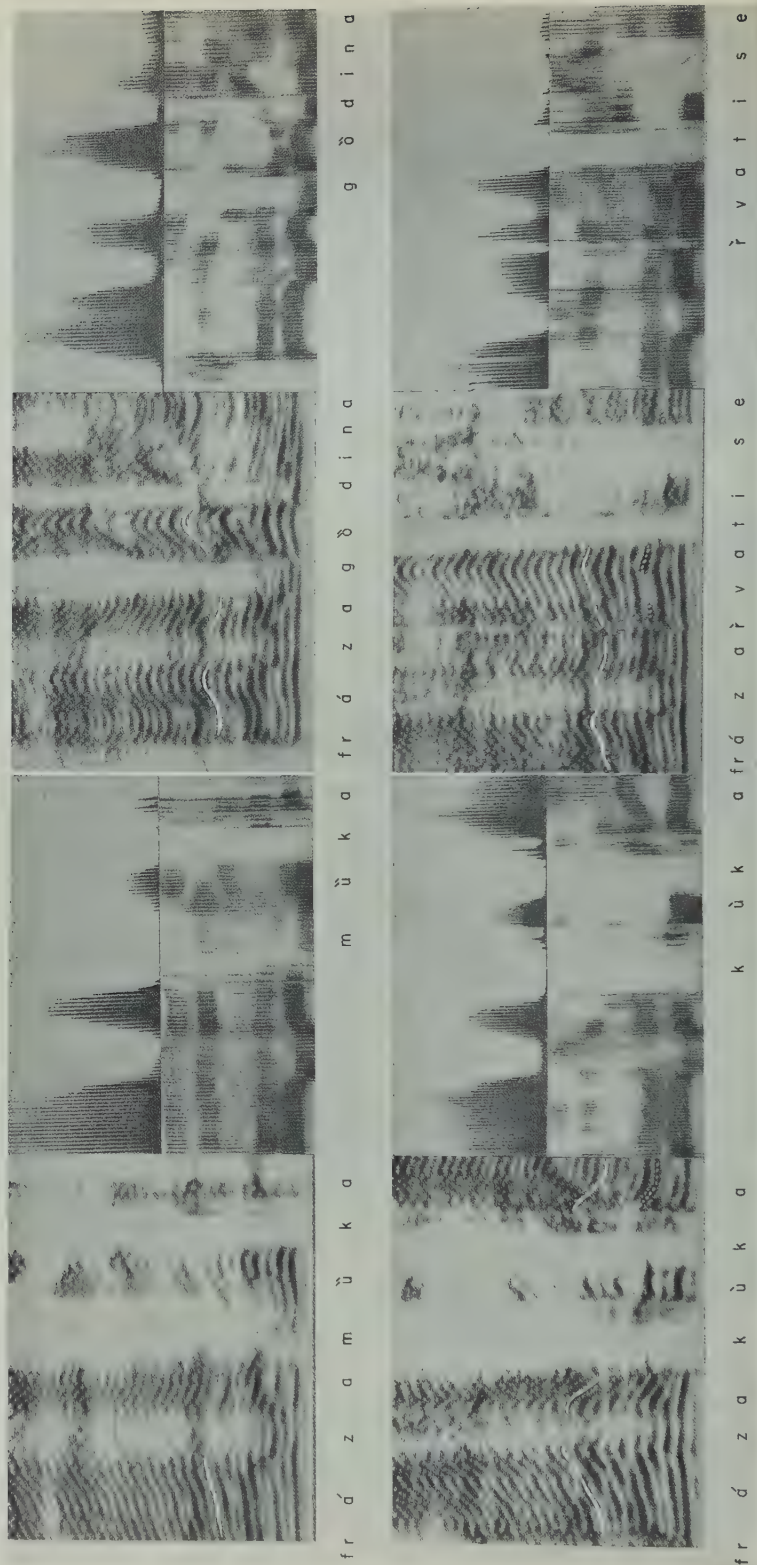


Fig. 2. Narrow-band spectrograms, broad-band spectrograms, and continuous amplitude displays of four utterances by informant DK.

continuous amplitude displays of four utterances by informant DK are presented. The fundamental frequency contours on the first syllables of the two words *mùka* and *kùka* are identical, except for the difference caused by the different initial consonants. The second syllable of *mùka* contains a laryngealized vowel, and none of the harmonics could be traced; the fundamental frequency is falling on the second syllable of *kùka*. The continuous amplitude display indicates that there is a considerable difference in the acoustic intensity of the /a/ in the second syllables of these two words; this difference will be discussed in a later section. The trisyllabic word *gòdina*, with a short falling accent on the first syllable, has a rising-falling fundamental frequency movement on the stressed syllable and falling fundamental frequency on the second and third syllables. The fundamental frequency on the phrase *rvati se*, with a short rising accent on the first syllable, is rising through the first syllable and the first part of the second syllable; after a peak in the second syllable, the fundamental frequency falls during the rest of the phrase.

### 3.3. Intensity patterns

Table II contains intensity measurements taken from the test words. The values were read from the output of a speech power measuring circuit, which was displayed on a Sanborn two-channel graphic recorder, and are expressed in decibels on an arbitrary scale reading from 0 to 50 db. The intensity values were measured at the beginning of the stressed vowel, at the point where peak intensity was reached, and at the end of the stressed vowel. The peak intensities on second and third syllables were also measured. Intensity patterns on words containing /r/ as syllable nucleus are presented separately (in table V); however, the second and third syllables of such words are also included in table II. Although the fundamental frequencies of laryngealized vowels were not included in the frequency table, the acoustic intensity of such vowels was averaged over the different peaks corresponding to the slow flaps of the vocal folds, and the corresponding measurements are included in the intensity table. A separate tabulation has been made of the amplitudes of second syllables with /a/ as the syllable nucleus. Direct comparison between the intensities of the second syllables of words with different accentual patterns is thus possible.

The intensity patterns associated with long falling and long rising accents will be considered first. The intensity pattern on the



Table II

Intensity Patterns in Test Words (Intensity in decibels)

Test word type	First syllable					Second syllable				Third syllable	
	No.	Beg.	Peak	End	Peak occurs at % of total duration	All vowels		/a/		No.	Peak
Short falling (29)											
Monosyllabic	11	28.0	42.0	33.0	58%						
	(-3r)										
Disyllabic	14	23.7	40.8	29.5	46%	15	32.7	10	34.8		
	(-1r)										
Trisyllabic	4	29.8	43.3	31.8	50%	4	37.3			4	31.0
Long falling (41)											
Monosyllabic	21	28.0	40.1	25.5	27%						
	(-5r)										
Disyllabic	20	20.9	41.2	28.1	25%	20	33.5	14	34.4		
	(-1r)										
Short rising (13)											
Disyllabic	8	27.8	41.7	32.0	44%	8	39.4	5	39.6		
	(-1r)										
Trisyllabic	5	33.0	42.0	35.0	58%	5	40.6	3	41.3	5	27.6
	(-4r)										
Long rising (17)											
Disyllabic	8	28.3	40.7	32.0	65%	8	38.7	7	39.1		
Single peak											
Disyllabic	(1. pk.	7	40.9		28%						
Two peaks	valley	21.1	38.9	32.2	55%	7	39.0	7	39.0		
	2. pk.)		41.0		91%						
Disyllabic	1r					1	39.0	1	39.0		
Trisyllabic	1r					1	37.0	1	37.0	1	32.0
100 occurrences of <i>fráza</i>											
Single peak	54	27.3	41.4	35.6	74%	54	37.3	54	37.3		
Two peaks	(1. pk.	46	41.1		31%						
	valley	28.3	39.5	34.9	62%	46	37.7	46	37.7		
	2. pk.)		41.2		83%						

words with long falling accent was characterized by a single intensity peak (in 95% of the test words) occurring at approximately  $\frac{1}{4}$  of the duration of the vowel, or in other words, coinciding with the pitch peak. The drop in intensity from the peak to the end of the stressed vowel was approximately 15 db. The second syllable had an amplitude that was approximately 7 db lower than the peak occurring on the first syllable. While the highest intensity peak always occurred near the beginning of the vowel, four different patterns of intensity distribution over the total duration of the vowel were

observed. Two of these occurred with equal frequency and constituted together 60% of the cases. In both patterns the syllable nucleus appeared to consist of two segments. In one case, the first part of the vowel was produced at a steady intensity level for about 34% of its duration, followed by a rapid decrease taking place during the production of the second half of the vowel. In the other case, two distinct intensity levels could be observed. In words exhibiting this type of intensity pattern, the first level had about 37% of the total duration, the second level approximately 22%. The difference between the two intensity levels was approximately 4 db; the fall from the end of the second level to the end of the vowel was approximately 10 db.

In 25% of the utterances, no segmentation into two parts appeared possible. The vowels had a smooth intensity curve, with a peak in the first quarter of the duration of the vowel.

In 5% of the utterances, two distinct intensity peaks were observed. The first peak occurred at approximately 20% of the duration, the second at 60% of the duration. The valley between the two peaks occurred at 40% of the duration, and was approximately 3 db lower in intensity than the first peak. The difference in intensity between the two peaks was approximately 0.5 db, the first peak having the greater intensity.

The remainder of the long falling patterns involved /r/ as syllable nucleus. Their intensity patterns will be described in a separate section.

Two different types of intensity patterns were observed on words with long rising accent. In approximately half of the instances, only one peak was observed; in the other half, two peaks appeared on the vowel with long rising accent. Of the 15 test words (2 words with /r/ as syllable nucleus are not included), 8 had a single peak on the first syllable, and 7 had two peaks. Of the 100 occurrences of the word *fráza*, 54 had a single peak, and 46 had two peaks. In the words with a single peak, this peak occurred considerably before the peak in fundamental frequency was reached (at 65% of the total duration vs. 97% for the fundamental frequency in the test words, and at 74% vs. 91% in the 54 occurrences of *fráza* with a single peak). In words with two peaks, the second peak occurred at approximately the same time when the fundamental frequency peak was reached. In disyllabic words with a single peak, the intensity drop from the peak to the end of the syllable was approximately 7 db; the second

syllable had an intensity that was 2–3 db lower than the peak occurring on the first syllable. In disyllabic words where two intensity peaks occurred during the first syllable, the first peak was reached at approximately  $\frac{1}{3}$  of the total duration, the second peak at approximately  $\frac{9}{10}$  of the duration, with a valley at approximately  $\frac{3}{5}$  of the duration of the vowel. On the average, the second peak was a fraction higher than the first peak (0.1 db); the valley between the two peaks had approximately 2 db less intensity than the peaks. The second syllable was approximately 2–3 db less intense than the first syllable in both types of words. In one occurrence of a trisyllabic word, the third syllable was 5 db weaker than the second syllable.

The words with short falling accent reached their intensity peak approximately in the middle of the word; the intensity dropped by about 10 db from the peak to the end of the stressed vowel. In disyllabic words, the second syllable had an intensity that was comparable to the intensity on the second syllable of words with long falling patterns; the amplitude was approximately 8 db lower than the amplitude of the peak on the first syllable. In the four trisyllabic words, the second syllable had approximately 6 db less amplitude, which then dropped by an additional 6 db on the third syllable.

The words with short rising accent reached their intensity peak at the middle of the accented syllable, being in this respect similar to the words with short falling accent. It should be recalled here that the intensity peak did not coincide with the peak in fundamental frequency in these words. The drop from the peak to the end of the stressed vowel was comparable to that on words with the short falling pattern (approximately 10 db). The amplitude on the second syllable, however, was noticeably different from that of the words with short falling accent. In disyllabic words with short rising accent, the peak on the second syllable was only approximately 2 db lower than that on the first syllable; in disyllabic words with short falling accent, the peak on the second syllable was about 8 db lower than the peak on the first syllable, resulting in a difference of approximately 6 db between the two types of disyllabic words<sup>14</sup>. On words with three syllables, the second syllable was only about 1 db less intense than the first, but the difference between the second and the third syllable was approximately 12 db.

Figure 3 contains reproductions of graphic displays of the intensity patterns of six utterances by informant DK. On each re-



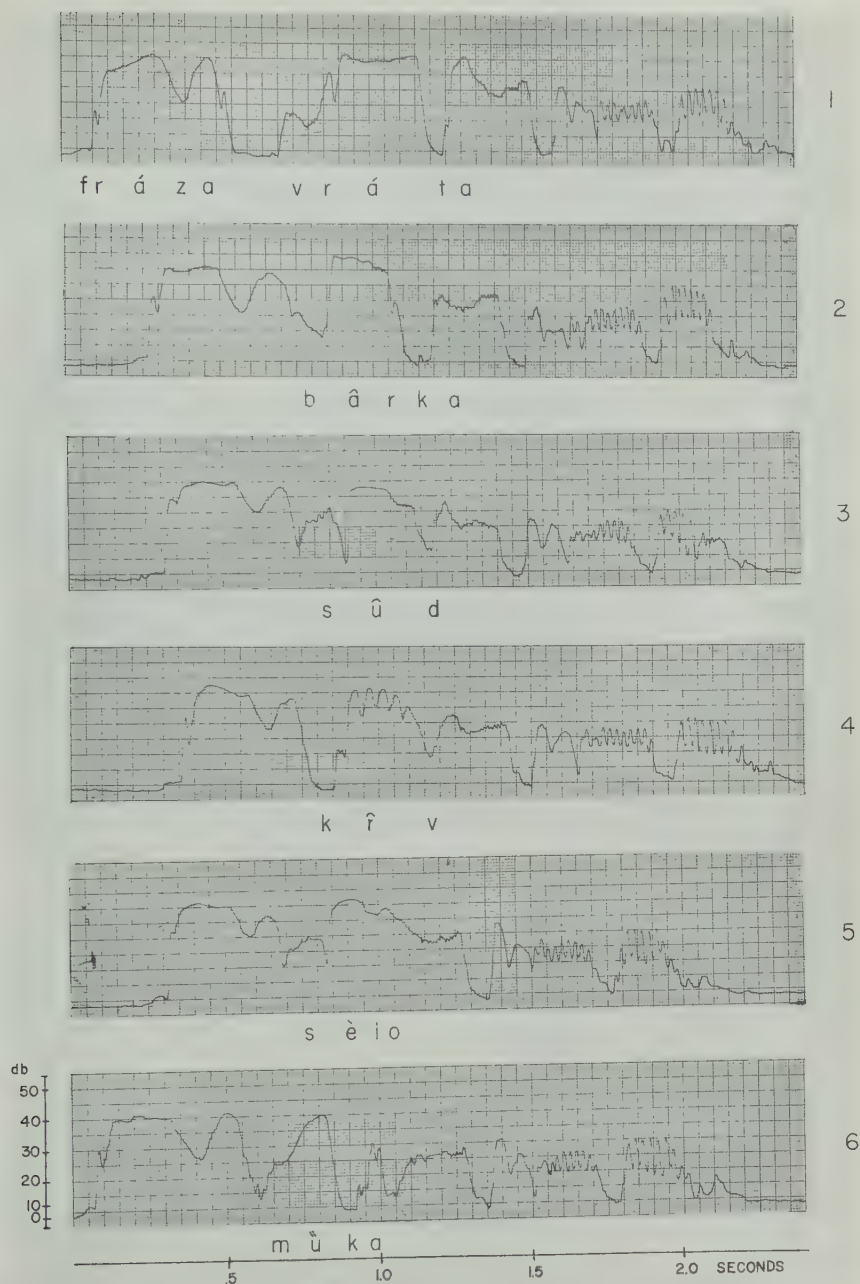


Fig.3. Reproductions of graphic displays of the intensity patterns of six utterances by informant DK. Intensity in decibels is indicated on the ordinate; time in seconds is indicated on the abscissa.

production, the ordinate of the trace represents the short-time average of the absolute value of the acoustic pressure wave (cf. reference<sup>11</sup>), expressed in decibels on an arbitrary scale. The circuit was practically linear over a range of 30 db. Time in seconds is indicated on the abscissa. On the original display, 1 cm = 10 db on the intensity scale, and 10 cm = 1 sec. on the time scale.

The six utterances contain the test words *vráta*, *bárka*, *súd*, *křv*, *sělo*, and *mùka* in the frame *fráza . . . je nàpīsānā na tábli*. The first test word, *vráta*, had long rising accent; the peak acoustic intensity of the vowels of both syllables was identical. The second sentence reproduced here contained the test word *bárka*, with the same vowels in both syllables as in *vráta*, but with long falling accent. The intensity pattern on *bárka* was characterized by falling intensity on the first syllable, and by considerably weaker intensity on the second syllable (the actual difference between the peaks in this utterance was 12 db). Utterance 3, containing the word *súd*, illustrates the intensity pattern characterized by two distinct levels which was observed in approximately one third of the words with long falling accent. Utterance 4 shows the intensity pattern on the word *křv*, illustrating the manifestation of the long falling accent on a monosyllabic word with /r/ as the syllable nucleus. The peaks correspond to the periods of voicing between the flaps constituting the trilled /r/. Utterance 5 contains the test word *sělo*, with short rising accent and relatively greater amplitude on the second syllable than in the word *mùka*, with short falling accent, represented in utterance 6. Six manifestations of the long rising accent may be observed on the repetitions of the word *fráza* at the beginning of each utterance. The speaker used very low pitch toward the end of each utterance; the fluctuations in the trace correspond to pressure changes associated with each vocal fold vibration.

In summary, we found that both the long accents had a set of phonetic characteristics that distinguished them clearly from each other and also from the short accents. The situation was ambiguous with regard to the two short accents. A difference in fundamental frequency patterns was observed only on words of three syllables; the two-syllable words, however, showed a difference in amplitude that could conceivably be sufficient to distinguish between the two types of word.

### 3.4. *The functional significance of the difference between the two short accents*

Some information about the relative importance of the above-described phonetic clues may be gained from the reactions of the listeners. As was mentioned at the beginning of this paper, the recorded tapes were played to informants DK and VJ, and their identifications and comments were noted. The two long accents were always clearly distinguished both from each other and also from the short accents. Ambiguities were noted in the identification of lexical items with the short accents. If only one form was possible, the identification was always positive. For example, the word *bòba* was identified as g. sing. of *bòb* "bean". In the standard language, *bòba* should have a short rising accent. We may thus assume that the speaker had intended to produce the form with the short rising accent. On the other hand, the word *kùka* was identified by both listeners as having two meanings, "hook" and "of the hip". When the tape was recorded, informant DK was requested to produce the g. sing. form of the word *kùk* "hip", which should have the form *kùka*, with short rising accent. The word *kùka* "hook", which was offered by the listeners as alternative meaning, has a short falling accent in the standard language. Similar ambiguities were observed in the other disyllabic words in which a difference in accent is expected to be associated with a difference in meaning.

The results of the listening test suggest that although a phonetic difference may be observed between the disyllabic words classified as belonging to the short falling and the short rising category, these average phonetic differences are not accompanied by a difference in meaning. This is a puzzling result, and there appears to be no immediately obvious explanation. A separate experiment should be designed to check the generality of the results.

One factor that corroborates the validity of the observed intensity differences is the fact that a similar difference was observed in the second syllables of disyllabic words with long rising and long falling accents. In this case there was also a clear difference in the fundamental frequency patterns in addition to the intensity differences in the second syllable.

### 4. *Other phonetic observations*

A number of other phonetic observations were made during the investigation, which will be briefly reported here.





#### 4.1. Accent and phonetic quality

The problem was investigated whether there was any connection between the selection of a particular allophone of the vowel and the accent associated with it. The positions of the first three formants of the stressed syllable nuclei were measured in the 100 test words and in the 100 occurrences of *fráza*. Table III presents the average frequency positions of the first three formants for each vowel, separated according to accent type. Table IV shows the variations possible for 100 repetitions of the same vowel with the same accent (/á/), and illustrates the ranges within which the averages may be expected to fluctuate. An inspection of table III indicates that accent has probably no influence on the phonetic nature of the vowel allophone. On the other hand, it appears clearly that the vowels /e/ and /o/ have two different allophones depending on the length of the vowel. This may be seen even more clearly from the acoustical vowel diagram presented in figure 4. Each point represents the target position of the stressed vowel in one of the 100 test words. The first formant frequency is indicated on the vertical axis, the second formant frequency on the horizontal axis. The values for /r/ were measured during the vocalic part between the flaps characterizing /r/; the vocalic periods obviously represent a central vowel. The wide dispersion of the points for /r/, compared, for example, with the compact distribution of the points for /a/, suggests that formant structure is probably not the essential acoustic characteristic of syllabic /r/.

Table IV

Formant Positions of /á/ in 100 Occurrences of *fráza* (Frequencies in cycles per second)

F <sub>1</sub>		F <sub>2</sub>		F <sub>3</sub>	
No. of occurrences	Frequency range	No. of occurrences	Frequency range	No. of occurrences	Frequency range
6	625-650	5	1200-1225	1	2300
91	675-700	39	1250-1275	16	2500
3	725-750	49	1300-1325	21	2525-2550
		7	1350-1375	39	2575-2600
				16	2625-2650
				7	2675-2700
100	625-750	100	1200-1375	100	2300-2700
Average	695		1285		2580

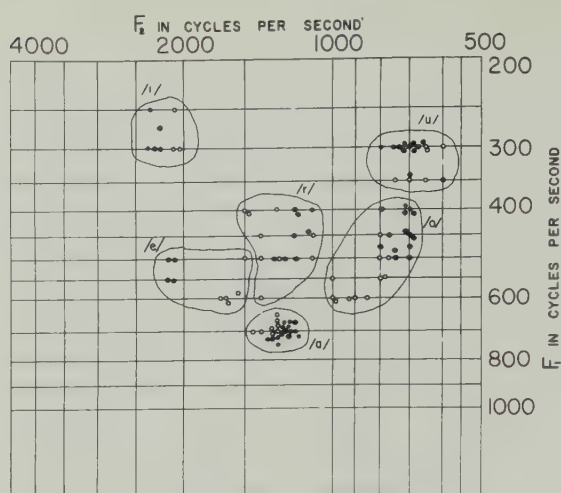


Fig. 4. Acoustical vowel diagram, representing the formant positions of the stressed syllable nuclei in 100 test words, uttered by informant DK. Circles indicate  $F_1$ - $F_2$  positions of short vowels, filled dots represent long vowels.

Duration appears to have no significant influence on the formant structure of allophones of /i/, /u/, /a/, and /r/. The vowels /e/ and /o/ have allophones [ɛ] and [ɔ] with short quantity, [e] and [o] with long quantity. Within the allophones determined by duration, accent appears to have no further influence.

#### 4.2. Duration

The contrast between long and short vowels was maintained in the stressed syllables. The average duration of the vowel of the first syllable in words with short falling accent was 12.1 centiseconds, in words with short rising accent 11.0 csec, long falling accent 25.2 csec, and long rising accent 26.4 csec. Thus the long syllable nuclei were somewhat more than twice as long as the short stressed syllable nuclei.

In unstressed position (i.e. in the second and third syllables of the test words) the durational contrasts appeared not to be maintained. Actual measurements were not made, since no segmentation point could be established between the end of the last syllable of the test word and the following word of the frame (the word *je* begins with a semivowel, so that formant movements are continuous). However, when the informants listened to the tapes, their comments

suggested similar confusions in identification as had been observed with the short falling and rising accents. For example, DK had been requested to produce the words meaning "of the king" and "of the kings", which are given as *králja* and *kráľja* in the standard language. The words occurred at different points in the random list, and so the informant was not consciously contrasting one with the other. Both words were identified as "of the king", or gen. sing. (i.e. with short vowel in the second syllable) by both listeners.

*Belic* has stated<sup>15</sup> that the contrast between short rising and short falling accents is preserved in disyllabic words with a long second syllable. The informants whose speech was investigated in the course of this study appeared to have no quantity distinctions in unstressed syllables. Since this distinction is not used by the informants, it appears entirely consistent that they also seemed to have no accent contrasts in disyllabic words, regardless of the length of the vowel of the second syllable in the standard language.

#### 4.3. *Influence of sentence intonation*

In several previous studies attempts have been made to distinguish the accents occurring on the separate lexical items from the superimposed sentence intonation. In his description of the manifestations of the various tones, *Appel*<sup>16</sup> indicated each time whether the word in question occurred in a rising or falling portion of the intonation contour. More recently, *Hodge*<sup>17</sup> has tried to solve the problem by describing both tone and intonation with four pitch phonemes and three degrees of stress. *Ivić*<sup>18</sup> offers an interesting suggestion that might help explain the loss of quantity distinctions in the unstressed syllables. According to *Ivić*, the apparent tendency in the language to eliminate from the end of the word distinctive elements such as quantity and accent might be explained as an attempt to eliminate the collision between word intonation and sentence intonation. A nucleus is formed in the first part of the word, which contains the distinctive "signs" of word intonation; the second part of the word, particularly the last syllable, thus becomes free to carry the functions of sentence intonation.

The present investigation confirms *Ivić*'s theory. In the materials recorded by informant DmK, there were several minimal pairs, produced in such a manner that a long rising accent could be observed with falling intonation, and a long falling accent with rising intonation. (The 100 frame utterances by DK all had an intonation



peak on the test word, and falling intonation on the second syllable of the test word.) In those instances, the influence of the intonation appeared on the second syllable. The pattern occurring on the first syllable was the same under both falling and rising intonations.

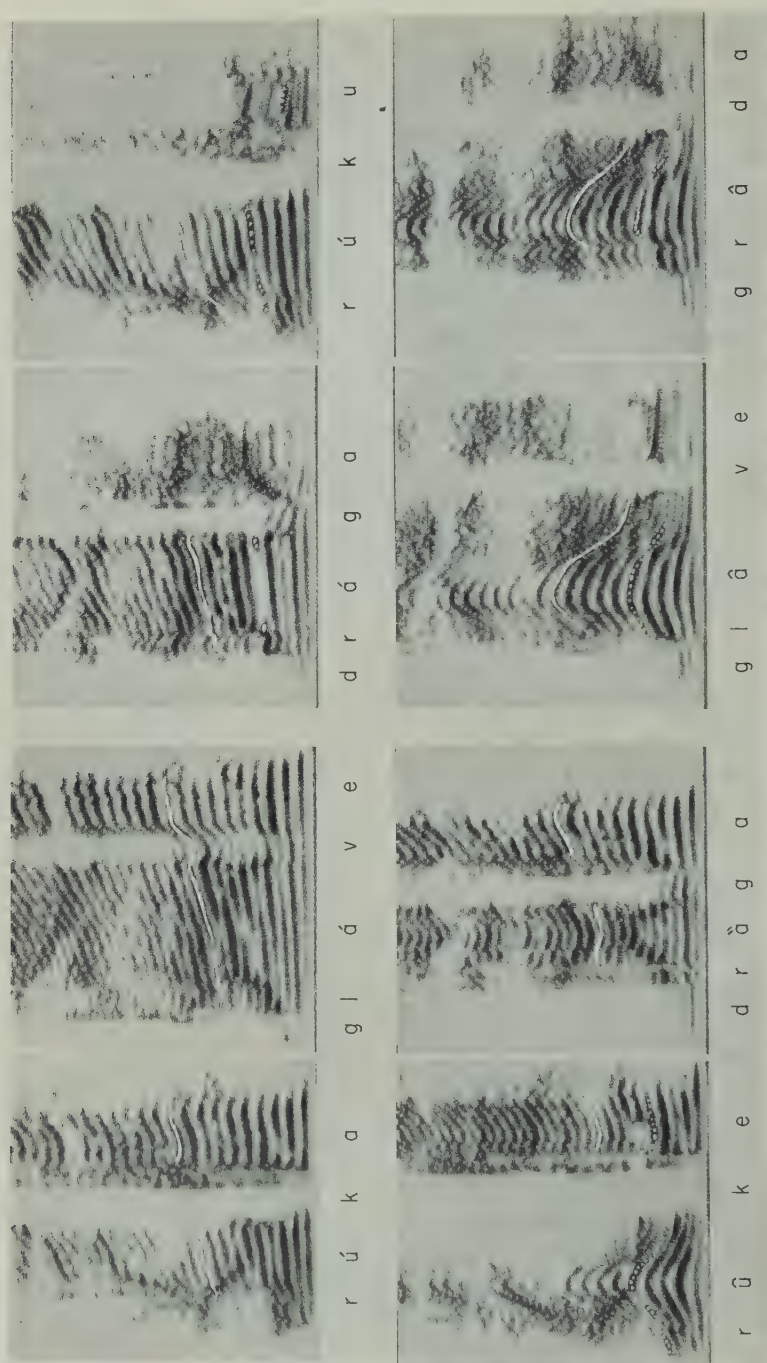
Figure 5 contains some examples illustrating the effect of intonation on word accent. Narrow-band spectrograms of eight words uttered by informant DmK are presented. Of the four words with long rising accent, *rúka* and *gláve* were produced with rising intonation, *drága* and *rúkū* with falling intonation<sup>19</sup>. It is evident from the fundamental frequency patterns that the effect of intonation was essentially confined to the second syllable. Three words with long falling accent and one word with short falling accent are included in the bottom half of the figure. Of these, the words *rúke* and *drāga* were produced with rising intonation, *gláve* and *grāda* with falling intonation. Again, it appears that in the interaction between word accent and sentence intonation, the fundamental frequency pattern on the first syllable signals the accent associated with the word, the fundamental frequency pattern on the second syllable carries information relative to sentence intonation.

#### 4.4 *The manifestation of the different accents on words with /r/ as syllable nucleus*

If either a pitch or intensity pattern is significant in distinguishing between words, it is interesting to observe how these distinctions are realized in words where the distinctive accent appears on /r/. The syllable nucleus /r/ in Serbo-Croatian is a tongue-tip trill, consisting of one flap (or at the most two) for the short syllabic /r/ and three to five flaps for the long syllable nuclei /ř/ and /ř̃/. The set of test words contained 17 occurrences of /r/ with all four accents; the data are presented in table V.

The measurement of the fundamental frequency during the periods of voicing between the flaps presented a particularly difficult problem. As far as this measurement could be carried through, it appears that the fundamental frequency patterns on the stressed syllables with /r/ as syllable nucleus were similar to those occurring on other syllable nuclei. The four trisyllabic words with short rising accent offer particularly good examples of the high fundamental frequency on the second syllable that has been described above as a feature of trisyllabic words with short rising accent. On the other hand, the one disyllabic word with short rising accent (*křsta*) had a

Word and accent type	Duration of /r/ in csec	Fundamental frequency in cps			3rd			Intensity in db		
		First syllable			First syllable			(Refer to text for column headings)		
Beg.	Peak	End	Beg.	End	Beg.	End	Beg.	End	2nd Peak	3rd Peak
Short falling										
vřt	11	110	128	120					25 43 30 38 30	
křst	10		130	126					23 44 32	
přst	9		123	120					5 44 32 39 30	
křvāv	14	120	130	120	122	97			38 29 39	41
Average	11	115	128	122	122	97			18 42 31 39 30	41
Long falling										
břk	26	120	125	75					30 43 32 43 36 42 36 39 33 36 12	
mřc	26	115	122	75					36 40 33 41 34 39 32 36 31 36 21	
křv	28	120	130	85					22 41 32 41 35 41 33 39 30 36 25	
cřn	24		130	100					19 41 32 40 35 41 36 39 31 33 32	
cřv	24		125	80					17 40 32 40 37 38 20	34
břka	18	120	125	92	80				23 42 30 42 30 40 10	
Average	24	122	126	85	80				24 41 32 41 34 40 28 38 32 35 22	34
Short rising										
křsta	9		126		lar.				15 43 28	34
řvati se	4	110	120		125	130	125	105	36 43 41	43 34
křstiti	8		120		135	120	100		10 44 35	40 32
třkati	12	125	132		145	126	115	100	20 41 30 38 18	41 34
grgljati	11		110		120	110	110	85	20 38 27 39 31	40 15
Average	9	118	122		131	122	113	97	20 42 32 39 25	40 29
Long rising										
Křsta	24	110	125		125	100			15 39 34 40 38 40 38 39 25	39
křcati	20	110	120		130	110	110	100	18 39 31 40 37 40 34	37 32
Average	22	110	123		128	105	110	100	17 39 33 40 38 40 36 39 25	38 32



*Fig. 5. Narrow-band spectrograms of eight words uttered by informant DmK.*

laryngealized second syllable and low intensity in the second syllable, which indicates that it was pronounced with phonetic features more often associated with the short falling category. The single occurrence of short falling accent on a disyllabic word with /r/ as syllable nucleus (*křvāv*) had, in turn, characteristics that are usually the property of words with short rising accent, having a second syllable with high intensity and a falling rather than low level fundamental frequency. This apparent contradiction is one more illustration of the observation that although certain phonetic features are present in the language that might potentially distinguish between the two short accents, they are not used in a systematic fashion to produce differences in meaning.

The intensities of the second and third syllables, apart from the contradiction that occurred in short disyllabic words, were similar to those observed in words with other syllable nuclei: second syllables of words with rising accents on the first syllable had higher intensities than second syllables of words with falling accents. The acoustic intensity patterns of the stressed syllables with /r/ as syllable nucleus are also contained in table V. The intensities between and during each tongue-tip closure were measured; the numbers in the corresponding column in table V refer to these values. An example of the type of trace from which the measurements were made (the intensity curve of the word *křv*) was included in figure 3.

There appears to be no observable difference between the intensities of the flaps or the voiced periods between them, when /r/ is produced with either the long falling or the long rising accent. This suggests that the intensity patterns on the stressed syllables carry less information than the fundamental frequency patterns, since the patterns on /r/ are phonemically equivalent to those occurring on other vowels, where phonetic differences in the intensity distributions were observed. The detailed shape of the intensity contours is thus probably non-significant; the major burden for differentiation appears to be carried by the quantity difference on the one hand, and fundamental frequency difference on the other hand.

### 5. Summary and conclusions

The phonetic characteristics of the long rising and long falling accents in the dialect of Belgrade have been described. In the material analyzed during the study, the long rising accent was



characterized by rising fundamental frequency and increasing intensity on the first syllable, followed by a second syllable with only slightly weaker intensity. The fundamental frequency on the second syllable started high; the drop in frequency occurred during the second syllable. The long falling accent was characterized by a decreasing fundamental frequency on the first syllable, accompanied by a decrease in intensity. The second syllable had considerably lower intensity. The fundamental frequency on the second syllable was usually level at the low end of the range of the speaker's voice. Sentence intonation affected the second syllable; therefore the fundamental frequency patterns on the first syllable appear to be distinctive.

There appeared to be no difference in the fundamental frequency patterns associated with disyllabic words with a short vowel in the stressed first syllable. In trisyllabic words with a short rising accent on the first syllable, the second syllable was found to be higher in frequency than the first syllable. The intensity patterns associated with the two types of words showed a difference in the intensity of the second syllable: the average for words where a short rising accent would be expected was considerably higher than the average for words where a short falling accent is expected. The difference was comparable to the intensity difference observed in the second syllables of two-syllable words with long rising and long falling accents. The listeners' reactions suggested that this phonetic difference was not accompanied by a difference in meaning, and thus should be considered phonemically nondistinctive.

The dialect under consideration is that of the younger generation of the educated circles of Belgrade. From the limited data available for the present study a tendency is noted in this dialect toward merging the two short accents. The phonetic features that characterize the two short accents appear to be losing gradually their distinctive function<sup>20</sup>. No phonetic differences in pitch are present in disyllabic words. An intensity difference is phonetically present in the majority of words, but nondistinctive. Trisyllabic words maintain both fundamental frequency and intensity differences at the present stage. The material available for the present study does not allow an estimate of the extent to which these conditions prevail in the usage of other speakers of the dialect. The dialect itself appears to be in a period of rapid change. As we are dealing with a changing system, it appears futile to try to set up a

static pattern to explain the function of the emerging features. Certain of the described elements probably constitute survivals of an older system; others indicate the direction of the change in progress. The present study is offered as an attempt to document one stage of this development.

### *Akustische Korrelate des Akzents im Serbokroatischen*

#### *Zusammenfassung*

Die phonetischen Merkmale der langen steigenden und fallenden Akzente im Belgrader Dialekt des Serbokroatischen wurden beschrieben. Zweisilbige Wörter mit langem steigendem Akzent hatten steigende Grundfrequenz und steigende Intensität auf der ersten Silbe und nur wenig schwächere Intensität auf der zweiten Silbe. Das Absinken der Grundfrequenz fand während der zweiten Silbe statt. Zweisilbige Wörter mit langem fallendem Akzent hatten sowohl fallende Grundfrequenz wie auch fallende Intensität auf der ersten Silbe. Die Intensität der zweiten Silbe war beträchtlich niedriger. Die Grundfrequenz der zweiten Silbe lag gewöhnlich unverändert auf der tiefsten Stufe des Stimmumfangs der Sprecher. Die Satzintonation beeinflusste die zweite Silbe; deshalb sind die Akzentphänomene der ersten Silbe als distinktiv anzusehen.

Im untersuchten Dialekt wurden keine Unterschiede im Tonverlauf der zweisilbigen Wörter mit kurzen Vokalen in der ersten betonten Silbe beobachtet. In dreisilbigen Wörtern mit einem kurzen steigenden Akzent auf der ersten Silbe hatte die zweite Silbe höhere Frequenz als die erste. Die Intensität der zweiten Silbe der Wörter mit kurzem steigendem Akzent war durchschnittlich beträchtlich höher als die Intensität der zweiten Silbe der Wörter mit kurzem fallendem Akzent. Der Unterschied ist vergleichbar mit dem Unterschied zwischen den zweiten Silben in Wörtern mit langem steigendem und langem fallendem Akzent. Hörerurteile bezeugen aber, daß dieser phonetische Unterschied nicht mit einem Bedeutungsunterschied assoziiert ist; deshalb muß der Intensitätsunterschied in der zweiten Silbe als nicht distinktiv angesehen werden. Es scheint also, daß die phonetischen Merkmale, die die beiden kurzen Akzente charakterisieren, im Begriffe sind, ihre distinktive Funktion zu verlieren. Im Tonverlauf wurden keine phonetischen Unterschiede bei zweisilbigen Wörtern beobachtet. Ein Intensitätsunterschied konnte phonetisch festgestellt werden, hatte aber keine phonemische Funktion. In dreisilbigen Wörtern wurden sowohl Tonverlaufs- wie auch Intensitätsunterschiede festgestellt. Die beiden kurzen Akzente sind also in zweisilbigen Wörtern zusammengefallen, werden aber in dreisilbigen Wörtern auseinandergehalten. Eine Spezialuntersuchung der beiden kurzen Akzente im Belgrader Dialekt wird vorgeschlagen.

### *Quelques corrélats acoustiques de l'accent serbocroate*

#### *Résumé*

On décrit les caractéristiques phonétiques des accents longs (croissants et décroissants) dans le dialecte serbocroate de Belgrade. Paroles à deux syllabes avec accent long et croissant ont une fréquence fondamentale et une intensité croissantes dans la première syllabe, et une intensité inférieure, tant soit peu, dans la deuxième syllabe. La décroissance de la fréquence fondamentale a lieu dans la deuxième syllabe. Paroles à deux syllabes avec accent long et décroissant ont une fréquence fondamentale aussi bien qu'une intensité décroissantes dans la première syllabe. L'intensité de la deuxième syllabe est considérablement inférieure. La fréquence fondamentale de la deuxième syllabe persiste, d'ordinaire, invariable sur le niveau le plus bas de l'étendue normale de voix du parlant.

L'intonation de la phrase influe sur la deuxième syllabe; c'est pourquoi il faut regarder comme distinctifs les phénomènes accentuels de la première syllabe.

On n'a pas pu découvrir, dans le dialecte examiné, des différences d'accent musical dans les paroles à deux syllabes avec voyelles brèves dans la première syllabe accentuée. Dans les paroles à trois syllabes avec un accent bref et croissant sur la première syllabe, la deuxième syllabe montre une fréquence supérieure à celle de la première. Dans les paroles avec accent bref et croissant, l'intensité de la deuxième syllabe est, en moyenne, considérablement supérieure à celle de la deuxième syllabe dans les paroles avec accent bref et décroissant. Cette différence est comparable à la différence qui existe dans la deuxième syllabe des paroles avec accent long et croissant d'une part, et accent long et décroissant d'autre part. Pourtant, selon le témoignage des écoutants, cette différence phonétique ne s'associe pas à une différence de signification, c'est pourquoi il faut regarder comme non-distinctive la variation d'intensité dans la deuxième syllabe. Il paraît donc que les caractéristiques phonétiques qui marquent les deux accents brefs sont en train de perdre leur fonction distinctive. Dans les paroles à deux syllabes on n'a pas pu observer des différences phonétiques d'intonation du mot. On a pu, pourtant, distinguer une différence phonétique d'intensité, mais qui n'avait pas de fonction phonémique. Dans les paroles à trois syllabes on a découvert des différences d'intonation aussi bien que d'intensité. Cela signifie que les deux accents brefs se sont confondus dans les paroles à deux syllabes, tandis qu'ils retiennent leur distinction dans les paroles à trois syllabes. On recommande une enquête spéciale pour les deux accents brefs dans le dialecte de Belgrade.

#### *References and Footnotes*

<sup>1</sup> *A. Meillet and A. Vaillant*, *Grammaire de la langue serbo-croate*, p. 28, Paris 1952.

<sup>2</sup> *Aleksandar Belić*, *Savremeni srpskohrvatski književni jezik*, p. 117, Beograd 1951.

<sup>3</sup> *Karl-Heinz Pollok*, "Zur Geschichte der Erforschung des serbokroatischen Akzentsystems", *Die Welt der Slaven* II, pp. 267–292 (1957).

<sup>4</sup> *W. Appel*, "Gestaltstudien, A. Untersuchungen über den Akzent in der serbokroatischen Sprache". *Wiener slavistisches Jahrbuch* I, pp. 53–70 (1950). The material analyzed by *Appel* consists of 17 utterances (10 words, 7 short sentences); the fundamental frequency patterns are obtained by means of a laborious procedure, which evidently placed an upper limit on the number of items that could be analyzed by this method. It is interesting to notice that, according to *Appel*, his data were used by *Trubetzkoy* when the latter formulated his theory about the Serbo-Croatian accents. Cf. *N. S. Trubetzkoy*, *Grundzüge der Phonologie*, 2nd ed., pp. 193–195, Göttingen 1958.

<sup>5</sup> *Djordje Kostić*, "O jačini naglaske dvosložnih reči pod (") i (') akcentom". *Južnoslovenski filolog* 18: 123–131 (1949–50).

<sup>6</sup> *Helmut Lüdtke*, "Das prosodische System des Urslavischen und



seine Weiterentwicklung im Serbokroatischen". *Phonetica*, Suppl. ad Vol. 4 (1959), p. 133: "Der Ausspracheunterschied besteht darin, daß im Falle *lúka-sèla* beide Silben des Wortes ungefähr gleich stark oder zumindest gleich hoch ausgesprochen werden, während im Falle *Lúka-sèla* die erste Wortsilbe deutlich stärker und höher ist als die zweite. Wenn außerdem der Ablauf des /*ū*/ in *Lúka* und *lúka* verschieden ist, so ist das eine ebenso irrelevante Begleiterscheinung wie die größere Intensität des /*l*/ in *sèla* gegenüber dem in *sèla*; entscheidend ist das Intensitätsverhältnis zwischen der ersten und der zweiten Silbe."

<sup>7</sup> *Milan Rešetar*, Elementargrammatik der serbokroatischen Sprache, 3rd ed. (by *Edmund Schneeweis*), p. 17, Max Niemeyer Verlag, Halle a.d. Saale 1957: "Der Akzent ist entweder fallend oder steigend. Bei dem ersteren ist der höchste Ton am Anfang der betonten Silbe, von dem er bis zum Ende der Silbe oder des Wortes fällt; bei dem letzteren steigt der Ton vom Anfang der betonten Silbe bis zum Ende und sinkt in der folgenden Silbe wieder ab."

<sup>8</sup> *A. Meillet* and *A. Vaillant*, op. cit., p. 22: "Une voyelle accentuée est prononcée à la fois sur une note plus haute et avec plus d'intensité qu'une voyelle inaccentuée. Mais l'intensité est ordinairement faible en serbo-croate; elle ne fait que donner à la syllabe accentuée un certain relief par rapport aux syllabes voisines, les syllabes inaccentuées gardant toujours leur netteté à côté des syllabes accentuées. L'élément musical (de hauteur) est prépondérant, et donne au serbo-croate une allure 'chantée' caractéristique."

<sup>9</sup> *Aleksandar Belić*, op. cit., pp. 115–116 (summarized by *I. L.*): "With the falling accents the expiratory force or stress and the musical accent or pitch are both highest at the beginning of the syllable, and the pitch drop is paralleled by a decrease in intensity towards the end of the syllable. In the first part of the syllable the expiratory force falls very abruptly, but in the second part it falls either slowly or not at all. The pitch drop is from a musical fifth to an octave in long falling accents, a musical second or third in short falling accents. With the rising accents, the musical tone rises, whereas the expiratory force increases during the first half of the accented syllable and decreases during the second half. The rise in pitch is approximately equal to a musical fifth in the long rising accent, a fifth or a third in the short rising accent. In certain dialects the rise in the fundamental frequency of the stressed syllable may be carried over into the next syllable, so that the beginning of this



syllable may have higher pitch than the stressed syllable. The expiratory stress remains on the accented syllable.”

<sup>10</sup> I would like to use this opportunity to thank the informants for their cooperation. Mr. *Dmitrije Kostić* made the recording during his visit to the Communication Sciences Laboratory of the University of Michigan in Ann Arbor in 1956. Both Mr. *Dušan Kalić* and Mr. *Vladimir Jovanović* gave freely of their time during their stay in Ann Arbor in 1960. I am particularly indebted to Mr. *Kalić* for the engineer's patience with a linguist's curiosity.

<sup>11</sup> *G. E. Peterson* and *N. P. McKinney*, “Measurement of speech power”. *Phonetica* 7: 65 (1961).

<sup>12</sup> This applies, of course, only to the falling pattern. The observation was made also by *Appel* (op. cit., p. 60). Similar segmental conditioning of the distribution of the fundamental frequency contour was observed on American English utterances where test words were pronounced with falling intonation (*G. E. Peterson* and *I. Lehiste*, “Duration of syllable nuclei in English”, *J. acoust. Soc. Amer.* 32: 693–703, 1960). Thus we appear to be dealing with a general phonetic phenomenon rather than with a feature characterizing the Serbo-Croatian accentual system.

<sup>13</sup> The main sources for this information were the handbook by *M. Rešetar* quoted in reference 7, and the dictionary of *S. Ristić* and *Ž. Kangrga* (*Svetomir Ristić* and *Jovan Kangrga*, *Rečnik srpskohrvatskog i nemačkog jezika*, Beograd 1928).

<sup>14</sup> A similar result was obtained by *Kostić* in the research described in reference 5.

<sup>15</sup> *A. Belić*, op. cit., p. 117.

<sup>16</sup> *W. Appel*, op. cit., p. 60.

<sup>17</sup> *Carlton T. Hodge*, “Serbo-Croatian Stress and Pitch”, *General Linguistics* 3, pp. 43–54 (1958).

<sup>18</sup> *Pavle Ivić*, “Die Hierarchie der prosodischen Phänomene im serbokroatischen Sprachraum”. *Phonetica* 3: 36 (1959).

<sup>19</sup> The word was produced by DmK with long rising accent and identified by the two other informants as the genitive plural of the word *rūka*. According to *Rešetar*, op. cit., p. 32, the g. pl. of this word should have the form *rūkū*. It was interesting to observe that informant DK used a similar substitution in giving the Serbian equivalent of the English phrase “of the legs”. The form used by DK was *nógu* rather than the expected *nǝgū* with short falling accent.

<sup>20</sup> Ten years ago, *Belić* (cf. reference<sup>15</sup>) had observed that the

distinction was lost in two-syllable words, excepting those that had a long unstressed vowel in the second syllable. The informants whose speech was analyzed in the course of this study have lost both the contrast between long and short vowels in the unstressed second syllable of disyllabic words, and the contrast between short rising and short falling accent in disyllabic words.

### *Selected Bibliography*

- Appel, W.*: Gestaltstudien. A. Untersuchungen über den Akzent in der serbokroatischen Sprache. Wien. slav. Jb. 1: 53-70 (1950).  
*Belić, A.*: Savremeni srpskohrvatski književni jezik. Prvi deo: glasovi i akcenat. Beograd 1951.  
*Hodge, C. T.*: Serbo-croatian stress and pitch. *General Linguistics* 3: 43-54 (1958).  
*Ivić, P.*: Die Hierarchie der prosodischen Phänomene im serbokroatischen Sprachraum. *Phonetica* 3: 23-38 (1959). - Die serbokroatischen Dialekte. (Mouton, 's-Gravenhage 1958).  
*Kostić, D.*: O jačini naglaske dvosložnih reči pod (") i (') akcentom. *Južnoslovenski Filolog* 18: 123-131 (1949-50).  
*Lüdtkje, H.*: Das prosodische System des Urslavischen und seine Weiterentwicklung im Serbokroatischen. *Phonetica* 4: 125-156 (1959).  
*Meillet, A. et Vaillant, A.*: La langue serbo-croate (H. Champion, Paris 1952).  
*Pollok, K. H.*: Zur Geschichte der Erforschung des serbokroatischen Akzentsystems. *Welt Slaven* 2: 267-292 (1957).  
*Rešetar, M.*: Elementargrammatik der serbokroatischen Sprache. 3rd ed. (by E. Schnee-weis), (Max Niemeyer, Halle-Saale 1957).  
*Ristić, S. i Kangrga, J.*: Rečnik srpskohrvatskog i nemačkog jezika. Drugi deo: srpsko-hrvatsko-nemački (Verlagsbuchhandlung Rajković und Čuković, Beograd 1928).  
*Trager, G. L.*: Serbo-croatian accents and quantities. *Language* 16: 29-32 (1940).  
*Trubetzkoy, N. S.*: Grundzüge der Phonologie. Prague, 1939. 2. Aufl. (Vandenhoeck und Ruprecht, Göttingen 1958).

### *Appendix*

This appendix contains the 100 test words analyzed in the course of the study. The words are presented according to accent type. Within each group, the words are arranged according to syllable nuclei in the sequence /i e a o u r/.

#### *1. Short falling*

Word	Form	Meaning and comments
dīm	n. sg. m.	smoke
līpa	n. sg. f.	lindentree
pīta	n. sg. f.	pie
hlēb	n. sg. m.	bread

Word	Form	Meaning and comments
sèla	n. pl. n.	villages
čas	n. sg. m.	hour
stati	inf.	to stand
bläto	n. sg. n.	swamp
vräna	n. sg. f.	crow
kräva	n. sg. f.	cow
jägoda	n. sg. f.	strawberry
järica	n. sg. f.	young goat
böga	g. sg. m.	God
pölje	n. sg. n.	field
göra	n. sg. f.	mountain
kös	n. sg. m.	woven basket
böb	n. sg. m.	bean
könj	n. sg. m.	horse
gödina	n. sg. f.	year
gödinä	g. pl. f.	of the years
rüm	n. sg. m.	rum
kük	n. sg. m.	hip
müka	n. sg. f.	trouble
müha	n. sg. f.	fly
küpim	l. sg. praes. ind.	I gather
vrt	n. sg. m.	garden
krist	n. sg. m.	cross
přst	n. sg. m.	finger
křvāv	n. sg. m.	bloody

## 2. Long falling

Word	Form	Meaning and comments
sîn	n. sg. m.	son
križ	n. sg. m.	cross
lipä	g. pl. f.	of lindentrees
pivo	n. sg. n.	beer
mêd	n. sg. m.	honey
lêp	n. sg. m.	beautiful
sêlā	g. pl. n.	of villages
krälj	n. sg. m.	king
grād	n. sg. m.	city

Word	Form	Meaning and comments
grādā	g. pl. m.	of cities (So DK; informant VJ said the word was in g.sg., which according to Rešetar, p.24, should be <i>grāda</i> , with short /a/ in the second syllable.)
prāvda	n. sg. f.	justice
krāvā	g. pl. f.	of cows
stāri	n. sg. m.	old
glāvu	acc. sg. f.	head
strāža	n. sg. f.	watch
zlāto	n. sg. n.	gold
bārka	n. sg. f.	boat
bôg	n. sg. m.	God
rôd	n. sg. m.	gender, kind
nôs	n. sg. m.	nose
bôj	n. sg. m.	battle
rôg	n. sg. m.	horn
gôrā	g. pl. f.	of mountains (this word occurred twice)
póljā	g. pl. n.	of fields (This word was produced w. a long falling accent by DK and accepted as such by VJ. Acc. to Rešetar, p.28, it should be <i>póljā</i> , with long rising accent.)
vôdā	3.sg.praes.ind.	he leads (So DK; VJ suggested, as alternative meaning, g.pl.f. "of water". According to Rešetar, p.31, g.pl. of <i>vôda</i> should be <i>vôdā</i> .)
kljúč	n. sg. m.	key
zúb	n. sg. m.	tooth
kût	n. sg. m.	angle, corner
sûd	n. sg. m.	court of justice
sûh	n. sg. m.	dry



Word	Form	Meaning and comments
rûgā	3. sg. praes. ind.	he mocks
kûpīm	1. sg. praes. ind.	I buy
gûrā	3. sg. praes. ind.	he pushes
Lûka	n. sg. m.	a man's name
bîk	n. sg. m.	mustache
mîc	n. sg. m.	carriage
kîv	n. sg. f.	blood
cîn	n. sg. m.	black
cîv	n. sg. m.	worm
bîka	g. sg. m.	mustache

### 3. *Short rising*

sèlo	n. sg. n.	village
sèla	g. sg. n.	village
žèna	n. sg. f.	woman, wife
nòga	n. sg. f.	leg
nògu	acc. sg. f.	leg
bòba	g. sg. m.	bean
jàrica	n. sg. f.	summer wheat
		(DK says that this word is not in his vocabulary. VJ said it was a repetition of the word <i>jàrica</i> "young goat".)
kùka	g. sg. m.	of the hip
		(Both informants said that this word could also mean "hook" in n.sg.f. According to the dictionary, the word for "hook" is <i>kùka</i> .)
křsta	g. sg. m.	cross
řvati se	inf.	to struggle, to wrestle
křstiti	inf.	to baptize
třkati	inf.	to race
		(According to the dictionary, the word should have a long rising accent: <i>třkati</i> .)
gřgljati	inf.	to gargle

## 4. Long rising

Word	Form	Meaning and comments
Míka	n. sg. m.	a man's name
žénā	g. pl. f.	of women
gláva	n. sg. f.	head
vráta	n. pl. n.	door
králja	g. sg. m.	of the king
kráľjā	g. pl. m.	of kings
ćása	n. sg. f.	little plate, round dish
nógu	g. pl. f.	of legs (According to Rešetar, p. 32, the form should be <i>nǝgǝ.</i> )
bója	n. sg. f.	color (This word was produced by DK with long rising accent. VJ identified it as g.pl.Acc. to the dictionary, n.sg. should be <i>bǝja</i> ; g.pl. should have the form <i>bǝjǝ</i> , acc. to Rešetar, p. 31.)
bǝjǝ	g. pl. f.	of colors
kǝnjǝ	g. pl. m.	of horses
vǝdǝ	g. pl. f.	of waters
lúka	n. sg. f.	harbor
súda	g. sg. m.	court of justice (this word occurred twice)
Krsta	n. sg. m.	a man's name
krcati	inf.	to load a ship (According to the diction- ary, the word should have short rising accent: <i>krcati.</i> )

Université de Nagoya

## Mémoire sur l'analyse de la qualité des voyelles

Par YOSHIYUKI OCHIAI

### *Explication des Termes Spéciaux*

- Confusion à l'Entrée – mode de confusion donnée par les chiffres dans des cellules en direction verticale, dans la matrice de confusion. Ce type de confusion est lié au mécanisme de l'établissement des qualités.
- Confusion à la Sortie – mode de confusion déterminée par les chiffres dans des cellules en direction horizontale, dans la matrice de confusion. Ce type de confusion est lié au mécanisme de la destruction des qualités.
- Importance Positive – distribution des qualités dans des régions de fréquences qui contribue positivement à l'établissement des qualités considérées.
- Importance Négative – distribution des qualités dans des régions de fréquences qui ne contribue pas à la formation des qualités, mais contribue plutôt à la destruction des qualités.
- Formant Positif – sommet de la structure formantique qui contribue subjectivement à l'identification et à la discrimination des phonèmes considérés, et qui est considéré objectivement comme se basant sur une action résonnentielle quelconque d'un circuit acoustique principal formé dans leur prononciation.
- Formant Négatif – foyer de vallon qui contribue subjectivement à l'identification et à la discrimination des phonèmes considérés (par exemple les phonèmes nasalisés), et qui est considéré objectivement comme se basant sur une action antirésonnentielle quelconque d'un circuit acoustique subsidiairement formé dans leur prononciation.
- Qualité Phonémique – une qualité du timbre de la parole, permettant l'identification et la discrimination des phonèmes au sens phonétique. Cette qualité est l'idée tirée du point de vue de netteté des paroles ou de l'intelligibilité du texte.
- Qualité Vocalique – une autre qualité du timbre de la parole, permettant l'identification et la discrimination des voix des parleurs. Cette qualité est l'idée tirée du point de vue de clarté des voix de ceux qui parlent.
- Répartition ou Distribution des Qualités – une idée sur le mode de distribution des qualités dans la dimension de fréquences. Plus précisément, la densité des qualités qui sont distribuées par unité des bandes de fréquences.
- Point de Balance en Qualité ou Point Bissecteur de Qualité – point d'intersection d'une paire de caractéristiques des qualités dans les distorsions des bandes-suppressions, de basse-coupures et de haute-coupures.
- Vallon Majeur – le plus grand vallon, c'est-à-dire la partie la plus vide dans le patron des voyelles. Cette position se situe actuellement et moyennement dans la région 1.3–1.6 kp/s au sujet du système des voyelles japonaises.

Matrice de Confusion – table sommaire des confusions représentée en forme carrée dont la ligne verticale donne les signaux et la ligne horizontale, les qualités correspondantes.

Comme nous l'avons déjà vu dans nos études<sup>6, 7</sup>, la qualité en général devient moindre quand la distorsion augmente : par exemple, dans le cas de la distorsion de bande-éliminatoire, on a la caractéristique de la qualité (tant de la qualité phonémique que de la qualité vocalique) qui diminue généralement avec l'agrandissement de la distorsion. La qualité diminue, parce que la bande transmissionnelle est limitée. Or, il y a deux sortes de limitation des bandes transmissionnelles : l'une est la limitation de bas en haut successivement en dimension de fréquences ; l'autre, celle de haut en bas, inversement. Nous avons ainsi une paire de caractéristiques de la qualité qui révèlent la phase de déperdition de la qualité, quand les signaux du timbre sont soumis aux distorsions réciproques des deux sortes. Cependant, c'est malheureux que les deux processus par deux distorsions nous amènent à des résultats un peu différents dans l'interprétation des deux caractéristiques ainsi obtenues. Par exemple, dans le cas de la qualité phonémique, la dérivée de la qualité par la fréquence n'est pas identique entre les deux distorsions ; l'une est un peu à côté de l'autre. La difficulté de l'interprétation du phénomène des qualités se révèle ici. Nous ne voulons pas aboutir à une interprétation déraisonnable.

Ce que nous pouvons incontestablement faire sur cette paire de caractéristiques des qualités (caractéristique en haute-coupure et caractéristique en basse-coupure), c'est de faire attention au point d'intersection de ces deux caractéristiques et de le considérer comme le point marquant qui divise la qualité entière en deux parties égales<sup>8</sup>. C'est ainsi que nous baptisons le point d'intersection «point de balance en qualité» ou «point bissecteur de qualité». Par l'introduction de ce point de balance dans la considération des qualités, nous pouvons mettre en lumière ces faits importants : pour la qualité phonémique, ce point de balance indique le point de concentration maximum de la qualité ; pour la qualité vocalique, ce point de balance indique, au contraire, le point de concentration minimum de la qualité. En d'autres termes, le point bissecteur en qualité phonémique des voyelles est à peu près le point indiquant le centre du domaine formantique qui contribue actuellement, parmi tous les pics, à la plus importante signification au sens phonémique ; le point bissecteur en qualité vocalique des voix, au contraire, coïncide



avec le point du vallon majeur (vocal glen, en anglais) qui est le point exact de la région d'amplitude presque minimum. Ces faits frappants semblent venus de la différence essentielle des modes de répartition des deux qualités. La qualité phonémique se répartit d'une manière très concentrée dans une région particulière relativement étroite. La qualité vocalique se répartit d'une manière dispersée et comparativement uniforme, s'étendant sur presque toute la région de fréquences. C'est la connaissance que nous pouvons avoir, grosso modo, sur la répartition des qualités. Nous devons en ajouter ici une autre chose sur la répartition plus détaillée des qualités, en nous appuyant sur nos expériences. La répartition de la qualité vocalique est divisée en deux parties, haute et basse, par le point du vallon majeur; la répartition de la qualité phonémique est un peu différente suivant la variation des voyelles. Des voyelles à formant simple et unique, comme les voyelles japonaises «A» et «O», ont habituellement une seule région de concentration de qualité\*. Quant aux voyelles à deux formants mal équilibrés, comme les voyelles «I» et «E», elles ont habituellement deux régions formantiques non balancées au sens de qualité phonémique: l'une à une concentration très faible, et l'autre, une concentration très forte. Il faut remarquer dans ce cas que la prédominance des régions au sens physique n'a pas nécessairement de relation avec la prépondérance des régions au sens subjectif. L'importance majeure, c'est-à-dire la prépondérance subjective au sens de qualité, ne correspond pas à la région de l'amplitude majeure au sens physique (région du formant majeur  $F_1$ ), mais elle correspond plutôt à la région de l'amplitude mineure (région du formant mineur  $F_2$ ); l'importance mineure, c'est-à-dire l'importance suivante, au sens subjectif, ne correspond pas à la région du formant mineur, mais à celle du formant majeur. Pour les voyelles à formant unique, le point de balance indique tout de suite le centre de l'unique concentration; et, pour les voyelles à deux formants, le point de balance tombe dans le voisinage de la région la plus concentrée. Quoi qu'il en soit, il est exact que le point de balance de qualité peut indiquer le point le plus marquant dans les patrons du timbre: ce point le plus marquant est celui de la concentration

---

\* Il faut noter dans ce cas que la position du pic maximum de l'unique formant ne correspond pas nécessairement à la place exacte de l'importance positive unique. Suivant nos expériences, il y a beaucoup de cas où elles ne coïncident pas exactement l'une avec l'autre. Ce fait signifie que le point du pic de ce qu'on appelle formant ne donne pas nécessairement la position la plus juste de la contribution maximum au sens phonémique.

maximum et non de la concentration minimum pour le patron phonémique, tandis que, pour le patron vocalique, c'est le point de concentration minimum et non celui de concentration maximum.

Comme l'inspection détaillée de la matrice de confusion composée par les signaux du timbre et les qualités de réponse le montre, les chiffres placés diagonalement dans cette matrice nous révèlent l'information sur le jugement juste dans nos perceptions des signaux; et les résultats du jugement erroné sont exprimés sommairement par les deux expressions, savoir par les chiffres donnés dans le côté perpendiculaire de cette matrice (confusion d'entrée) et par les chiffres donnés dans le côté horizontal (confusion de départ). Ces confusions de deux sortes, d'entrée et de départ, nous permettent de schématiser entièrement le phénomène de la qualité<sup>9</sup>.

En introduisant les idées sur la confusion à l'entrée ou à l'arrivée (*incoming confusion*, en anglais) ainsi que sur la confusion au départ ou à la sortie (*outgoing confusion*), nous pouvons absolument parfaire notre théorie de qualité. C'est parce que, par cela seulement, nous pouvons mettre en lumière non seulement le mécanisme de la déperdition de qualité, mais aussi celui de la formation de qualité<sup>10</sup>. En jugeant des caractéristiques de confusion, le cours de la formation de qualité est entièrement différent de celui de la déperdition de qualité. C'est ainsi que nous pouvons préparer une nouvelle voie dans l'interprétation du phénomène des qualités. La confusion au départ n'apporte rien de nouveau, parce qu'elle ne montre que le phénomène de déperdition des qualités que nous avons déjà pu suffisamment deviner par les allures générales des caractéristiques des qualités. Tandis que l'introduction de la confusion à l'arrivée nous amène à la découverte de l'idée de la formation des qualités que nous ne pouvons entrevoir par la seule étude des caractéristiques en général.

Revenons à notre thèse du point d'intersection. Le point d'intersection d'une paire de caractéristiques de la confusibilité au départ nous montre aussi ce que nous donne le point de balance déjà décrit. C'est ce qui est donné par le point d'intersection dans les deux caractéristiques de confusibilité à l'arrivée, que nous avons pu mettre en attente. Maintenant donc nous allons décrire plus en détail ce dont il s'agit quant au point d'intersection de la confusibilité à l'arrivée. Ce que signifie ce point est tout à fait différent selon la différence de qualité: pour la qualité phonémique, ce point d'intersection de la confusibilité à l'arrivée indique le centre de la

région de nuisance (ou d'importance négative) pour l'établissement de la qualité phonémique; pour la qualité vocalique, ce point montre, au contraire, le centre de la région d'importance (ou plus exactement d'importance positive) pour l'établissement de la qualité vocalique.

La parole est sans doute le phénomène humain en ce qui concerne l'expression et la transmission de l'intelligence et de l'émotion des humains. C'est sa forme acoustique qui rend service à la perception et la reconnaissance de la parole. Les paroles considérées dans le domaine de la communication sont donc un objectif physico-acoustique ainsi qu'un objectif psycho-physiologique. Il y a beaucoup de difficultés dans l'étude de la qualité des paroles. La plus grande existe dans la méthodologie par laquelle nous avançons dans nos études. Sur notre route, il y a des marécages infranchissables, des labyrinthes sans issues et des culs-de-sac qui ne conduisent nulle part. La mise en œuvre de la recherche sur les qualités, c'est exactement faire le pont entre la psychométrie et la physico-phonétique. Tout ce qui marche sur cette voie, doit éprouver, un jour ou l'autre, des difficultés pour se débarrasser du marais fangeux où il a mis les pieds ou pour trouver l'issue des impasses où il s'est égaré inconsciemment.

Le plus important et le plus vital dans la considération de qualité, c'est de ne pas confondre le phénomène physique avec le psychologique.

Suivant l'expression inverse, faire la distinction la plus claire entre le stimulus comme signal et la réponse comme qualité est tout à fait l'essentiel auquel nous devons tendre toujours. La confusion entre eux est la cause de tous les enchevêtrements dans les considérations.

Nous pouvons appeler, par exemple, la structure des cimes dans les patrons du timbre des vocales, «structure formantique» ou simplement «formant». C'est absolument la dénomination au point de vue physique. D'autre part, nous pouvons étudier les allures de la réponse correspondant au formant comme stimulus. Nous pouvons ainsi déterminer subjectivement ce qu'on appelle importance comme contribuant à l'établissement de la qualité. Nous avons alors le *formant* comme objectif physique d'une part, et d'autre part, l'*importance* comme objectif psychologique.

Comme principe général, l'importance d'une région de fréquences peut être définie par le degré de contribution de cette région



à l'établissement de la qualité considérée, et cette importance peut être actuellement représentée et mesurée par la valeur de la différence entre confusion à l'entrée et confusion au départ. Il y a naturellement deux sortes d'importance suivant le signe de cette valeur: l'importance positive qui correspond à la condition que la valeur de la confusion à l'entrée est plus grande que la confusion au départ; et l'importance négative qui correspond à la condition inverse où la confusion au départ est plus grande que la confusion à l'entrée.

Nous avons ainsi définitivement admis qu'il y a conséquemment deux sortes d'importances, positive et négative, si nous définissons l'importance comme donné ci-dessus. Cependant, quant au formant, y a-t-il aussi le formant positif ainsi que le formant négatif? Au sens général et assez large, nous devons peut-être admettre, au domaine de la qualité phonémique, l'existence du formant négatif: par exemple, comme nous l'avons déjà suggéré dans nos études sur les voyelles nasalisées<sup>11, 12</sup>, l'absence du formant particulier causée par l'absorption de la cavité buccale est indispensable pour la discrimination phonémique de certains sons nasaux. Ceci signifie qu'on doit admettre, dans les cas particuliers, que la prépondérance au sens négatif, c'est-à-dire la régression des amplitudes des composants de telle ou telle région, peut contribuer à l'établissement de la qualité phonémique.

Cependant, pour ce qui concerne les voyelles orales seulement, il n'est pas nécessaire d'insister sur l'existence du formant négatif. Par conséquent, faire la distinction entre «formant positif» et «formant négatif» est sans signification, au moins pour la présente discussion des voyelles orales.

Finalement, nous voulons discuter un peu concernant les conditions les plus essentielles qui déterminent et conditionnent la formation des voyelles orales. Pour caractériser et classer les voyelles, il y a sans doute une méthode qui s'appuie exclusivement sur les marques physico-acoustiques de leurs structures spectrales<sup>16, 1</sup>, et qui probablement est nécessaire et urgente pour le problème actuel de la discrimination automatique ou mécanique des sons de la parole. Il y a aussi une autre méthode qui est d'autant plus scientifiquement exacte qu'elle profite exclusivement du côté subjectif des paroles et souligne leur nature psychologique.

Dans la description ci-dessus, nous avons employé la dénomination des voyelles en nous basant simplement sur leurs formants: voyelle à formant unique et voyelle à deux formants. Sans doute,



cette dénomination est un peu conventionnelle, bien qu'elle soit très facile à comprendre. Pour l'exprimer avec une plus grande exactitude, nous devons nous attacher à l'*importance* au lieu du *formant*. Par exemple, ce n'est pas exact d'appeler la voyelle «A» «voyelle à formant unique», parce qu'il y a évidemment plusieurs formants dans le patron du timbre de cette voyelle. En toute rigueur, nous devons l'appeler «voyelle à importance unique». De même, au lieu de la dénomination «voyelle à deux formants», nous devons dire «voyelle à deux importances».

Il va sans dire que la voyelle à importance unique signifie plus rigoureusement que cette voyelle possède une seule et unique importance positive, et davantage qu'en dehors de cette importance positive, il y a deux importances négatives qui entourent l'importance première. De même, la voyelle à deux importances signifie que cette voyelle est caractérisée par deux importances positives et une importance négative intercalée entre les deux importances premières (voir la figure 1). Ces expressions-ci sont sans doute précises et générales. Les conditions données par les deux sortes d'importance, positive et négative, sont nécessaires et suffisantes. Donner l'importance positive seulement, c'est la condition nécessaire pour caractériser les

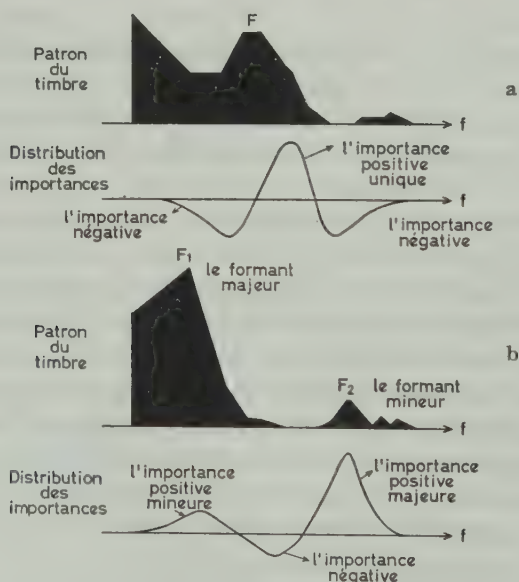


Fig. 1. La représentation des patrons du timbre et la répartition des importances.  
a: Voyelle à formant unique; b: Voyelle à deux formants non-balancés.

voyelles, mais ce n'est pas suffisant. On doit donner l'importance négative en dehors de l'importance positive pour avoir à la fois les conditions nécessaires et suffisantes.

En résumé, l'essentiel dans la considération de la qualité des voyelles et des paroles réside dans le dualisme de la qualité et non dans le monisme. Cette dualité de qualité est représentée par les qualités phonémique et vocalique suivant notre terminologie<sup>13,14,15</sup>. Le Prof. R. Husson emploie les termes «coloration vocalique» et «timbre extra-vocalique»<sup>2,3,4</sup>. Suivant sa terminologie, A. Moles utilise «personalisation» et «typification»<sup>5</sup>. Bien que les terminologies soient un peu différentes suivant les auteurs, il devient de plus en plus évident que l'on ne peut plus interpréter les structures formantiques des voyelles par l'aspect monotone et primitif de l'unicité de la qualité. Cela signifie que nous ne devons pas rester au point de vue primitif dans l'interprétation de l'information humaine. On doit avancer et adopter ici le dualisme: l'information intellectuelle et l'information émotionnelle. C'est tout ce qui est attribuable à la formation de la mentalité et à la nature de l'esprit humain.

Il est très difficile de donner la forme à la qualité, quelle que soit la signification de cette qualité. A plus forte raison, quand il s'agit de la qualité de la parole, parce que la parole est le sommet de toutes les choses que les humains ont jamais créés et produites. Donner une forme particulière à cette qualité et ensuite déterminer et mesurer sa grandeur, c'est une œuvre, sans exagérer, infiniment difficile.

### Résumé

Pour vérifier non seulement objectivement mais aussi subjectivement, le fait subsistant des deux qualités des voyelles, c'est-à-dire la coexistence de la *qualité phonémique* et de la *qualité vocalique*, sous l'aspect de la structure du timbre, nous avons fait une suite d'expériences objectives et subjectives; expériences de nature physique: obtenir le type (ou patron) phonémique et le type (ou patron) vocalique des voyelles soutenues; expériences de nature psychologique: mesurer à la fois la qualité phonémique et la qualité vocalique des voyelles. Pour la qualité du timbre, nous considérons exclusivement les signaux des voyelles en structure de fréquences ou en structure des harmoniques. Afin d'étudier la structure du timbre des voyelles en connexion avec leurs qualités, il vaut mieux introduire une idée de qualité, la répartition de la qualité dans la dimension de

fréquences, comme base de la théorie de qualité. C'est dans ce but que nous utilisons la distorsion de suppression des bandes. En examinant l'allure des caractéristiques des qualités dans les deux distorsions des suppressions, de bas en haut et de haut en bas, nous pouvons ainsi atteindre la notion de la distribution des qualités au long de l'échelle de fréquences. Pour étudier de plus près notre problème sur l'analyse des qualités, nous adoptons ici de préférence la méthode d'observation plus détaillée des confusions au sujet de la perception des voyelles déformées par la distorsion. Quant à la confusion du timbre, il y a essentiellement deux directions à considérer: *confusion à l'entrée* et *confusion à la sortie*. Pour la cause de la confusion qui paraît, nous considérons: Quand un son est mis en déformation par une distorsion quelconque, ce son devient très dissemblable du son original sans déformation (c'est-à-dire il sort du son original au sens de confusion) d'une part, et d'autre part, ce son devient semblable à un autre son quelconque (c'est-à-dire il entre dans un autre son au sens de confusion). Par cette observation complète de confusion, on peut atteindre cette fois les notions les plus importantes de la qualité qui s'établit ou de la qualité qui se détruit à cause de la distorsion. Voilà les notions d'*importance positive* et d'*importance négative* que nous présentons ici. Il en résulte que pour les phonèmes par exemple, il y a non seulement la région d'importances positives (ou *région utile*) mais aussi la région d'importances négatives (ou *région nuisible*). Quand une voyelle est prononcée convenablement, ses composants caractérisants viennent naturellement dans la région utile et propre. Mais quand une voyelle est mal prononcée ou quand elle est déformée au cours de sa transmission et que ses composants principaux et caractérisants viennent ou sont déplacés vers la région d'importance négative, il en résulte que cette voyelle est facilement confondue avec une autre voyelle dont la région utile coïncide accidentellement avec la région nuisible de la voyelle premièrement considérée.

### *Zur Analyse der Vokalqualitäten*

#### *Zusammenfassung*

Um die unter dem Aspekt der Struktur der Klangfarbe zu unterscheidenden zwei Vokalqualitäten – *phonemische* und *stimmliche* – sowohl objektiv als auch subjektiv zu verifizieren, wurde eine Reihe physikalischer und psychologischer Experimente durchgeführt. Bei den ersteren handelte es sich darum, den phonemischen und den stimmlichen Strukturtyp ausgehaltener Vokale zu bestimmen, bei den letzteren um die gleichzeitige Messung der phonemischen und der stimmlichen Qualität der Laute.

Hinsichtlich der Qualität der Klangfarbe betrachten wir ausschließlich die den Vokalen als Signal zuzuordnenden Frequenz- und Partialtonstrukturen. Um die Struktur der Klangfarbe im Zusammenhang damit zu untersuchen, empfiehlt es sich, als Grundlage der Theorie der Qualität den Gedanken einer Verteilung der Qualität in der Dimension der Frequenz einzuführen. Wir machen uns dabei die Verzerrungen zunutze, die aus der Beschneidung des Frequenzbereichs resultieren; untersucht man das Verhalten der aufsteigenden und absteigenden Qualitätskennlinien, so wird die Verteilung der Qualitäten über die Frequenzskala faßbar.

Eine eingehendere Untersuchung unseres Problems macht die sehr detaillierte Beobachtung des Verhörens erforderlich, wie es bei verzerrten Vokalen auftritt. Es sind hauptsächlich zwei Formen der Klangfarbenverwechslung zu unterscheiden: *initiales* und *finales* Verhören (confusion à l'entrée, confusion à la sortie).

Über die Ursachen des Verhörens lassen sich folgende Überlegungen anstellen. Wird ein Laut durch eine beliebige Verzerrung entstellt, so wird er einerseits dem unverzerrten Original sehr unähnlich (d. h. er geht im Prozeß des Verhörens aus dem ursprünglichen Laut hervor), andererseits einem anderen Laut sehr ähnlich (d. h. er geht im Prozeß des Verhörens in einen anderen Laut über). Eine vollständige Erfassung des Verhörens ergibt also die außerordentlich wichtigen Begriffe einer Qualität, die bei Bandbeschneidung gebildet, und einer Qualität, die bei Bandbeschneidung zerstört wird. Wir charakterisieren diesen Sachverhalt durch die Termini «positive Relevanz» (importance positive) und «negative Relevanz» (importance négative). Die Phoneme besitzen beispielsweise nicht nur einen Frequenzbereich positiver Relevanz («nützlicher Bereich»), sondern auch einen solchen negativer Relevanz («schädlicher Bereich»). Bei richtiger Aussprache eines Vokals fallen seine charakteristischen Komponenten natürlich in den nützlichen, d. h. adäquaten, Bereich. Wird ein Vokal dagegen nicht korrekt ausgesprochen oder bei der Übertragung deformiert, dergestalt, daß seine Hauptkomponenten und Charakteristika in Richtung auf den Bereich negativer Relevanz verschoben werden, so wird der Vokal leicht mit einem anderen verwechselt, dessen nützlicher Bereich zufällig mit dem schädlichen Bereich des ursprünglichen Vokals zusammenfällt.

### *Note on Analysis of Vowel Quality*

#### *Summary*

In order to verify, objectively and subjectively, the existence of two qualities of vowel sounds, namely, *phonemic quality* and *vocal quality*, from the standpoint of timbre structure, we have carried out the following series of studies: experiments of physical nature in which we have given the phonemic patterns as well as the vocal patterns of vowel sounds, experiments of psychological nature in which we have given the phonemic quality as well as the vocal quality. With regard to the timbre qualities, here only we consider the vowel signals in frequency structure or in harmonic construction. In the theoretical consideration of the timbre-signal construction in association with the timbre quality, it is most fitted to introduce the idea of the distribution of quality in the frequency dimension. To this end, we purposefully utilize the distortion of band-elimination type. Judging from the forms of the characteristics of qualities in the two distortions of band-elimination, from low to high and from high to low, that is, from the increasing and decreasing forms of quality characteristics in the distortions, we can very easily attain the conception of quality distribution in the frequency dimension. In order to study this problem more precisely and to make clear the nature of quality more specifically, we have adopted here the method of observing so-called confusions as explicating the details of quality phenomena in distortion. As to the timbre confusion,



there are two ways of observation: *incoming confusion* and *outgoing confusion*. For the cause of timbre-signal confusion, we consider the case where a certain signal is so deformed that it is easily confused with some other signal. This deformed signal has no more resemblance to the original signal without distortion (that is, this signal goes out of the original signal or it starts from the original signal, in the terms of confusion), but this signal rather resembles some other signal (that is, in the terms of confusion, it arrives in some other signal). By a complete observation and description of confusion phenomena, it is finally possible to discover and attain the most important idea of quality, that is, the idea of the quality which is formed and the idea of the quality which is destroyed respectively, resulting from the band-cutting distortions applied. We can express these two ideas by *positive importance* and *negative importance* respectively. As a general outcome of our experiments, it can be said that for phonemes, for example, there is not only the frequency-region of positive importance (that is, *useful region*) but also the frequency-region of negative importance (that is, *harmful region*). We explain more concretely. When a vowel signal is pronounced properly, its characteristic components do not fail to come exactly into the useful region of this signal, which enables us to identify this signal properly. But when this signal is pronounced incorrectly or when a correctly pronounced signal is deformed to some measure by some distortion in transmission, thus having its powerful components displaced towards the harmful region, this signal would easily be confused with some other signal, because the useful region of the latter signal coincides incidentally with the harmful region of the former.

### Bibliographie

1. *Fant, G.*: Acoustic analysis and synthesis of speech with applications to Swedish. Reprint from Ericsson Technics No. 1, 1959.
2. *Husson, R.*: Facteurs acoustiques des voyelles humaines (non nasalisées) et fréquence de coupure du pavillon pharyngobuccal. *Comptes Rendus*. t. 244: 1261-1263 (1957).
3. *Husson, R.*: Einige Ergebnisse von Untersuchungen der akustischen Struktur der Vokale. *Phonetica* 2: 193-198 (1958).
4. *Husson, R.*: La voix chantée (Gauthier-Villars, Paris 1960).
5. *Moles, A.*: Sur la «caractérisation» du discours et de la diction. *Cah. Acoust.* 76: 1-12 (1957).
6. *Ochiai, Y.*: Phoneme and voice identification studies using Japanese vowels. *Language and Speech* 2: 132-136 (1959).
7. *Ochiai, Y.*: Fondamentales des qualités phonémiques et vocaliques des paroles par rapport au timbre, obtenues en employant des voyelles japonaises vocalisées par des sujets japonais. *Mem. Fac. Engng., Nagoya* 10: 197-201 (1958).
8. *Ochiai, Y.* and *Fukumura, T.*: Timbre study of vocalic voices viewed from subjective phonal aspect; Part I. *Mem. Fac. Engng., Nagoya* 8: 77-94 (1956).
9. *Ochiai, Y.* and *Fukumura, T.*: Timbre study of vocalic voices viewed from phonal aspect; Part III. *Mem. Fac. Engng., Nagoya* 8: 222-239 (1956).
10. *Ochiai, Y.* and *Fukumura, T.*: Preliminaries to analysis of quality in speech communication. *Mem. Fac. Engng., Nagoya* 9: 306-315 (1957).
11. *Ochiai, Y.*: Memoirs on nasals. *Mem. Fac. Engng., Nagoya* 9: 147-153 (1957).
12. *Ochiai, Y.*; *Fukumura, T.* and *Nakatani, K.*: Timbre study on nasals; Part II. *Mem. Fac. Engng., Nagoya* 9: 160-173 (1957).
13. *Ochiai, Y.*: Mémoire sur les sons des voix humaines. *Mem. Fac. Engng., Nagoya* 4: 83-86 (1952).

14. *Ochiai, Y. and Yamashita, T.*: On timbre quality. Mem. Fac. Engng., Nagoya 7: 66-71 (1955).
15. *Ochiai, Y. and Oda, M.*: Sur l'intensité sonore subjective des voyales soutenues ayant les significations phonémique et vocalique. Mem. Fac. Engng., Nagoya 11: 103-111 (1959).
16. *Warschawskij, L. A. i Litwak, I. M.*: Issledovanie formantnovo sostava i nekotorykh drougikh fizicheskikh kharakteristik zvoukov rousskoy retchi. Problemuy fiziologit-cheskoy akoustiki, III 5-17. (La déduction de l'inventaire de formants et de quelques autres caractéristiques physiques des sons de la parole russe. Problèmes de l'acoustique physiologique, III, 1955.)

Adresse de l'auteur: Prof. Dr. Y. Ochiai, Université de Nagoya, Furo-cho, Chikusa-ku, *Nagoya* (Japon)

---

## Libri

---

### The Nagoya Group of Research on Speech Communication

#### A Review of Some of their Publications

A considerable number of research groups in Japan are currently undertaking work on problems of speech communication, but few of them report their results in Occidental languages. The group led by Professor *Toshiyuki Ochiai* at the Laboratory of Audiology, Nagoya University, Nagoya, however, does report its research progress in English, French and German. A selection of reprints of publications of members of the Nagoya group was sent to this Journal for review, and the undersigned, who is now at the Massachusetts Institute of Technology on leave from the University of Electro-Communications in Tokyo, was asked to be the reviewer. The titles of the papers received are:

*Y. Ochiai* and *H. Kato*: Sur la netteté et la naturalité de la voix humaine réfléchies du point de vue de la qualité de transmission. *Memoirs of the Faculty of Engineering, Nagoya University (M.F.E.)* 1: 109-115 (1949).

*Y. Ochiai*: Mémoire sur les sons des voix humaines. *M.F.E.* 4: 83-86 (1952).

*Y. Ochiai*: Memoirs on nasalics. *M.F.E.* 9: 147-153 (1957).

*Y. Ochiai*; *T. Fukumura* and *K. Nakatani*: Timbre study on nasalics, Part II: Preliminary experimental representation of timbre-patterns of sustained nasals. *M.F.E.* 9: 160-173 (1957).

*Y. Ochiai*: Fondamentales des qualités phonémique et vocalique des paroles par rapport au timbre, obtenues en employant des voyelles japonais vocalisées par des sujets japonais. *M.F.E.* 10: 197-201 (1958).

*T. Fukumura*: Quality problems in perception of repeated damped-sinusoids. *M.F.E.* 11: 112-120 (1959).

*Y. Ochiai* and *M. Oda*: Sur l'intensité sonore subjective des vocales soutenues ayant les significations phonémique et vocalique. *M.F.E.* 11: 103-111 (1959).

*Y. Ochiai*: Etude plus détaillée sur l'enveloppe des patrons du timbre des vocales orales, particulièrement au point de vue de la structure des vallons. *M.F.E.* 11: 89-102 (1959).

*Y. Ochiai*: Phoneme and voice identification studies using Japanese vowels. *Language and Speech* 2: 132-136 (1959).

*Y. Ochiai* and *T. Fukumura*: On the fundamental qualities in speech communications, principally on vowel quality. Technical Report of Research Developments at the Electric Communication Laboratory of Japan Telephone-Telegraph Co. 8: 1469-1549 (1959) (in Japanese).

Although the titles of the papers cover a variety of topics, it is clear that all of the investigations undertaken by Dr. *Ochiai*'s group were designed to answer one basic question: What are the "fundamental qualities" of speech sounds and what are their acoustic correlates? By "fundamental qualities" they mean not only the quality which determines phonemic identity, but also the qualities that constitute naturalness of speech sounds. As one important and measurable aspect of naturalness, they selected speaker identifiability, in the hope that study of the factors relating to the identification of the speaker by means of psycho-acoustic experiments would reveal the nature of voice quality on the acoustic level. From the point of view of communication engineering they assert, quite correctly, that besides phonemic identity, there are some other "qualities" that must be retained in the message. They propose that speaker identifiability is one of the most useful measures for evaluating communication channels, particularly those of high fidelity.

The experiments of the group have chiefly exploited two techniques: frequency analysis and listening tests. Analysis methods have been standard, including the use of the sonagraph in recent years. Vowel utterances of a limited number of speakers have been studied intensively and spectral peculiarities were examined for various conditions of vocalization in the course of searches for the acoustic correlates of voice quality as well as the cues for phoneme identification. Similar studies were made of nasal consonants and nasalized vowels with emphasis on the function of the nasal cavity as a determinant of voice quality. The psychoacoustic tests typically utilized vowels spoken in isolation and required the listeners to identify both the phoneme and the talker. In a series of studies systematic spectral distortion of the signals was introduced. The major objective of this psychoacoustic testing program was to observe the errors the subjects made. Confusion matrices of the subjects' responses are presented in detail and are analyzed to determine the contribution to the phoneme and speaker identification of various frequency components.

The detailed discussions of data in these papers contain many interesting suggestions. The experimental data they have provided constitute a useful source of information for the workers in this field. Some of their results, e.g., that the effect of spectral distortions on the accuracy of speaker identification is substantially different from that on vowel identification, are not of little significance.

The reports suffer, however, from a terminology that is misleading or vague and whose physical meaning is not clarified thoroughly. It sometimes is difficult to draw a clear-cut conclusion from their arguments, and some of their apparent conclusions may invite controversy. Ambiguity stems partly from the statement of the problem itself. Without defining the "fundamental qualities," either on the level of production or perception, it is hard for anyone to discuss, in any language, what their manifestations are in the spectral structure or what role they play in speaker identification.

Apparently, Professor *Ochiai* hoped somehow to find the solution of the most difficult basic problem in the patterning of the experimental results. The experiments were designed with a broad interest involving many aspects of the general problem rather than in such a particular manner that the results would either corroborate or refute a specific hypothesis or a model. While the former approach to many complex problems seems to be rather common in groups in many countries, the results of the present elaborate studies suggest that it is not a promising methodology for advancing fundamental knowledge in this field. The weakness of the approach is more apparent, but perhaps less harmful, when there is not an overpowering quantity of mechanically processed data.

The work of this group would have been much more illuminating, at least for some readers, if more confidence had been displayed in the acoustic theory of speech production. There is little doubt, with our present understanding of the speech production mechanism, that the phonetic value of the vowel, which can be described reasonably accurately by a set of about three lower formant frequencies, carries the principal information necessary for phoneme identification. But the phonetic value of a speech sound is not necessarily irrelevant to speaker identification. Even within a dialectically uniform community, there certainly is fluctuation which characterizes individual speakers. This characteristic of a speaker, however, is essentially different from what we would call his personal "voice quality," e.g., a "ringing" quality of his voice.

It would have been quite reasonable, therefore, if the phonetic value of the vowel had been separated from the rest of the spectral features as an independent dimension which can be described quantitatively. The so-called voice quality, as far as spectral aspects are concerned, can be defined rigorously as the remainder of the spectrum after subtracting the resonant characteristic that is determined, for example, by the best fitting formants with some standard damping. This measure, of course, is not always attributable only to the characteristics of the glottal source wave. It is possible that a



subsidiary articulatory feature (e.g., slight nasalization, or a peculiarity of glottal action) which is always associated with the production of some (or all) vowels of a particular speaker gives rise to a concomitant acoustic (gross) feature (e.g., highly damped formants). In discussing the nature of the subjectively perceived "voice quality" there is perhaps no reason to separate the acoustic consequence of the subsidiary articulatory features from the voice source characteristics – a point of view that the authors seem to take. The voice quality as tentatively defined above, could be evaluated in terms of such measures as high frequency-low frequency balance, statistically defined local irregularities, etc., and will primarily pertain to speaker identification. The contributions to the speaker identification of these features and of the phonetic value (if such contribution exists) may be estimated by the psychoacoustic techniques the Nagoya group has employed. It would be most desirable, however, if an indisputable proof can be supplied by data concerning subjects' responses to synthesized speech sounds, the pertinent acoustic features of which are precisely and selectively controlled.

The examination of speech sounds in the frequency domain, ignoring non-stationary characteristics in the time domain, may not enable us to determine all the acoustic correlates of voice quality. Among other effects, the contribution of a glottal fluctuation to speaker identifiability must be studied in the future.

It may be worth mentioning that the Nagoya group is one of the oldest speech groups in Japan; today there are many other groups with different inclinations and experimental techniques. In view of their respectable achievements and research potential, it is sincerely hoped that these groups will spend more time in communicating with western researchers. Sometimes the language barrier is not easy to overcome, but it is to be hoped that western readers will demonstrate understanding interest so that any sincere effort in this direction may be encouraged. The reviewer will be happy to supply any information which may be helpful for correspondence. As a general reference, it may be advised that the Journal of the Acoustical Society of Japan, published quarterly in Tokyo, occasionally carries interesting papers in this field, written in Japanese but with English abstracts.

Author's address: Prof. Osamu Fujimura, University of Electro-Communications, Chofu City, Tokyo (Japan).

**Aktuelle Probleme der Phoniatrie und Logopädie.** Vol. 1. S. Karger Basel/New York 1960. XVIII + 226 S., 74 Abb., sFr. 38.–.

Der vorliegende Band ist als Festgabe dem hochverdienten Zürcher Phoniater *Richard Luchsinger* zum 60. Geburtstag gewidmet. Der Herausgeber, *Felix Trojan*, Wien, umreißt im Vorwort die Zielsetzung des Werkes: die Mittelpunktstellung der Phoniatrie und Logopädie im Kreise einer Anzahl von Wissenschaften dadurch sinnfällig zu machen, daß jeder einzelne Aufsatz ein Thema behandelt, das die Sprach- und Stimmheilkunde jeweils zu einer Nachbarwissenschaft in Beziehung setzt.

Der internationale Rang des Werkes wird durch ein zweites und ein drittes Vorwort der derzeitigen Präsidentin der Internationalen Gesellschaft für Logopädie und Phoniatrie, *Joan Van Thal*, London, und *J. Tarneaud*, Paris, unterstrichen. Jedem der Beiträge, die in drei Sprachen vorliegen, sind Zusammenfassungen in deutscher, englischer und französischer Sprache beigelegt.

Der Lebenslauf und ein Verzeichnis der wissenschaftlichen Arbeiten von *R. Luchsinger*, das 100 Veröffentlichungen aufführt, lassen erkennen, daß der Jubilar heute als Forscher wie als Organisator zu einem Mittelpunkt der Welt seines Faches geworden ist.

Der Band gliedert sich in drei Abschnitte, in welchen die Verbindungen des Fach-

gebietes zu den Nachbardisziplinen unter drei Aspekten – A. dem physiologischen, B. dem physikalischen und C. dem psychologischen – behandelt werden. Von den 30 Beiträgen hervorragender Fachvertreter können bei der Mannigfaltigkeit der Themen in einem kurzen Referat nur einige Punkte herausgegriffen werden.

#### *A. Der physiologische Aspekt in der Phoniatrie und Logopädie*

Die ersten Beiträge sind Stellungnahmen zu Fragen aus dem Problemkreis der nervösen Steuerung der Stimmlippschwingung, wie sie anlässlich der Kontroverse um die These von *Husson* in den letzten Jahren in den Vordergrund der Diskussion getreten sind.

*Lullies* und *Paulsen* untersuchen die *afferenten Nervenimpulse*, welche von Rezeptoren der Muskulatur und besonders der Schleimhaut des Kehlkopfs ausgelöst werden und auf reflektorischem Wege die Stimmbandschwingungen beeinflussen. Die Bedeutung solcher Afferenzen für die Phonation besteht in der reflektorischen Unterstützung der Feineinstellung der Spannung und Länge der Kehlkopfmuskeln und damit auch der Regelung der Frequenz der Stimmbandschwingungen.

*Vallancien* kommt zu ähnlichen Schlußfolgerungen bei der Auswertung *röntgenkine-matographischer Ergebnisse*. Danach können die Vorgänge an der Glottis beim Gesang durch aerodynamische Gesetze ausreichend erklärt werden.

*Van den Berg* beschäftigt sich mit den *Registern der Singstimme* im Zusammenhang mit den Eigenschaften der Stimmbänder sowie des Aufhängesystems des Larynx und der aerodynamischen Verhältnisse. Er kommt zu dem Schluß, daß die Längsspannung der Stimmbänder sowie der Stimmmuskeln diejenigen Parameter sind, die die Hauptregister entscheidend beeinflussen.

Arbeiten aus dem Gebiet der *Genetik* und der *physiologischen Phonetik* schließen sich an.

*Pfändler* behandelt Ergebnisse der *phoniatischen Erbforschung* und belegt anhand eines neuen italienischen Stammbaumes die Möglichkeit eines gemeinsamen genetischen Ursprungs von Poltern und Stottern.

Die *Beziehungen des Vokalismus zu den Registern* betrachtet *F. Trojan* unter dem Gesichtswinkel der «Entwicklungsphonetik». Er deutet, in Fortführung seiner bereits 1950 aufgestellten These, den Vokalismus als Superstruktur auf der Basis der Register. Ein Orientierungsversuch an amphoteren Sprechτόnen (Sprechen von «I» und «U» sowohl mit vorwiegendem Brust- als auch mit Kopfregister) und seine röntgenologisch erhobenen Ergebnisse (Ausmaß der Lageveränderungen im Mundrachen und Kehlraum) werden mitgeteilt. Die Arbeit verdient das besondere Interesse des Stimmbildners, der erfahrungsgemäß schon immer zur Erzielung eines fehlenden Kopfklanges bei der Voix mixte zu Vorderzungenvokalen gegriffen hat, ohne daß bisher eine ausreichende theoretische und experimentelle Klärung dieser Zusammenhänge vorgelegen hätte.

*Physiopathologische und chirurgische* Beiträge schließen den ersten Abschnitt ab.

Frau *Mitrinowicz-Modrzejewska* zeigt anhand genauer klinischer Untersuchungen, daß man das *Gaumenspaltenleiden* heute nicht mehr als isoliert umschriebenen Defekt sehen darf, sondern im Rahmen einer gesamtkörperlichen Berücksichtigung als Systemleiden auffassen muß. Zahlreiche andere Abnormitäten der Spaltträgerpersönlichkeit lassen sich aufdecken. Der Referent möchte auf diese Arbeit besonders hinweisen, da eine ganzheitlich ausgerichtete logopädische Betreuung, wie wir sie heute anstreben, durch die vorgelegten Untersuchungsergebnisse auf eine breite Basis gestellt werden kann.

Mit der *chirurgischen Versorgung des Spaltträgers* befaßt sich *Croatto*, der über die Resultate seiner schon früher angegebenen Operationsmethode berichtet, welche durch Verdopplung der hinteren Pharynxwand eine Verbesserung des velo-pharyngealen Abschlusses und damit eine günstige Beeinflussung des offenen Naselns ermöglicht.

Spezielle Probleme der *Kehlkopfchirurgie* behandelt *Perello*; Fragen der sprachlichen Rehabilitation nach Zungenentfernung (*Glossektomie*) werden von *Brodnitz* erörtert.

#### *B. Der physikalische Aspekt in der Phoniatrie und Logopädie*

*Berendes* nimmt zur *Meßtechnik bei der Auswertung von Kehlkopf-Hochgeschwindigkeitsfilmen* kritisch Stellung. Er fordert sowohl Messung der Stimmlippenabstände von der Mittellinie als auch Flächenmessung des Glottisbildes (*Planimetrie*), da – insbesondere bei unregelmäßiger Gestalt der Glottis – die Anwendung nur einer Methode zu falschen Ergebnissen führen kann. Abstandsmessung und Flächenmessung müssen einander ergänzen.

*Svend Smith* berichtet über einen neuen Apparatetypus (Aerometer), der es erstmals erlaubt, die *Volumengeschwindigkeit der Luftströmung durch Mund und Nase* (Luftmenge in der Zeiteinheit) minutenlang ununterbrochen getrennt zu registrieren. Eine funktionelle Bestimmung der Nasenatmung ist dadurch möglich geworden.

Ein Verfahren zur *getrennten Aufzeichnung des nasalen und oralen Schalldruckes* gibt *Krech†* an. Es handelt sich um eine Weiterentwicklung der Methode *Doubek*. Der orale und der nasale Schalldruckanteil werden mittels eines Zweikanalmagnettongeräts über isolierte Kristallmikrophone durch einen Dämpfungsschreiber getrennt aufgezeichnet. Wenn das Verfahren auch keine unmittelbaren Aussagen über die Nasalität zuläßt, so dürften die gewonnenen Schalldruckdifferenzen in db doch dazu beitragen, sich dem Komplex der Nasalität von einer Teilkomponente her zu nähern, und darüber hinaus gibt es uns eine Möglichkeit an die Hand, phoniatische und logopädische Ergebnisse (z. B. Operationserfolge) zu normieren und exakt zu vergleichen.

*Winckel* betont in seinem Beitrag über die *Grenzanalyse des Komplexlautes*, daß für das Studium der Lautbildung die Spektralanalyse des quasistationären Lautzustandes zwar wertvolle Einblicke in die akustische Struktur der Einzellaute vermittelt hat, daß sie jedoch dem dynamischen Charakter der Stimme und Sprache im fließenden Lautstrom nicht ausreichend gerecht wird. Sowohl beim Sprechen als auch beim Singen finden ständig Veränderungen der Tonhöhe, der Lautstärke und – das erscheint besonders wichtig – auch der Formantfrequenzen statt. Die gegenseitige Beeinflussung benachbarter Laute im Lautstrom ist so beträchtlich, daß es beispielsweise nicht gelingt, aus einem besprochenen Tonband Vokale und Konsonanten herauszuschneiden und zu andern Wörtern zusammenzusetzen. Für gewisse Konsonanten ist gezeigt worden, daß sie gar keine selbständigen Gebilde sind, sondern wirkliche Übergänge (*transients*) darstellen. Unter diesen Aspekten müssen sowohl für die Analyse als auch für die Synthese die *Ausgleichsvorgänge* berücksichtigt werden und darüber hinaus die phonemischen bzw. sogar die Silbeneinheiten als maßgeblich für die physiologisch entsprechende Artikulationsdynamik angesehen werden.

*Tarnóczy* gibt phonetische Gesichtspunkte bekannt, die bei der Zusammenstellung von *Textlisten für Verständlichkeitsmessungen* zu berücksichtigen sind. Um die Kombinationsfähigkeit der Hörer, welche das Ergebnis einer Verständlichkeitsmessung nach der positiven Seite verfälschen könnte, auszuschließen, müssen die Verständlichkeitstexte zwar den Lautverbindungsgesetzen einer Sprache folgen, dürfen sich aber nicht nur auf die tatsächlichen, in der Wortbildung gebrauchten Möglichkeiten der Sprache beschränken. So empfiehlt *Tarnóczy* nebeneinander – bei gleicher Gesetzmäßigkeit – sinnvolle und sinnlose Lautfolgen, welche auf Grund einer statistischen Bearbeitung einer Einzelsprache zusammengestellt sind und die Lautverbindungsgesetze dieser Sprache widerspiegeln.

Von großer Bedeutung für den klinisch tätigen Laryngologen sind die *stroboskopischen Befunde bei frühen Stadien von Stimmbandkarzinom und Tuberkulose*. *Schönhärl*, dem wir eine inzwischen erschienene Monographie über die Stroboskopie verdanken, stellt hier folgenden Grundsatz auf: Nicht jeder stroboskopische Stillstand einer Stimmlippe muß



notwendig einen bösartigen Tumor oder eine Tuberkulose bedeuten, jedoch wird umgekehrt jeder Tumor und jeder Tuberkulosefall eine auffallende Bewegungseinschränkung bzw. einen Stillstand der Stimmlippenbewegung aufweisen. Daraus ergibt sich die Wichtigkeit der stroboskopischen Untersuchungsmethode für die Früherfassung gefährlicher Kehlkopferkrankungen.

### C. Der psychologische Aspekt in der Phoniatrie und Logopädie

Der dritte Teil des Werkes umfaßt zunächst Aufsätze aus den Grenzgebieten der Neurologie und Psychiatrie, um den engen Zusammenhang der seelischen Vorgänge mit dem Zentralnervensystem zu unterstreichen.

Er wird eingeleitet durch ein *phoniatisch-neurologisches Grundsatzreferat* von M. Seeman, in welchem die zahlreichen Querverbindungen zwischen den beiden Fächern angedeutet werden.

Wir wissen heute, daß dem Stottern – einem in sich sehr uneinheitlichen Symptomenkomplex – in vielen Fällen als Teilursache eine frühkindliche Hirnschädigung zugrunde liegt. R. und A. Schilling teilen eine Reihe von Spezialuntersuchungen und deren Ergebnisse mit, die geeignet sind, den neurologisch-phoniatrischen Befund zu ergänzen und die *Diagnose eines frühkindlichen Hirnschadens bei Stotterern* zu erhärten. Es handelt sich um das Elektroenzephalogramm, das Nystagmogramm, die Untersuchung der Feinmotorik nach dem Oseretzký-Test, die Messung der Vibrationsempfindung, die Röntgenzwerchfellkymographie und Lateralisationsproben.

Mit *Lateralisationsproben* (Untersuchung der Händigkeit), *Reifezeichen* und der *Elektroenzephalographie* beschäftigt sich auch die Arbeit von Subirana und Oller-Daurella. Die aufgezeigten Beziehungen der Reifezeichen zur Hemisphärendominanz dürfen als interessanter Beitrag zu dem immer noch weitgehend ungeklärten Fragenkomplex der pathologischen Linkshändigkeit und ihrer Zusammenhänge mit Sprachstörungen besonders hervorgehoben werden.

Weiss fand einen überraschend hohen Prozentsatz von *Stammelfehlern bei Schizophrenen*, vor allem Sigmatismen. Das Stottern war nicht häufiger als in der Durchschnittsbevölkerung.

Einen kasuistischen Beitrag zum autistischen Kindheitsmutismus verdanken wir Arnold. Unter dem Titel *«Schreiben statt Sprechen»* beschreibt er das Krankheitsbild eines vierjährigen Knaben, der unfähig war zu sprechen, der jedoch über eine ungewöhnliche Fertigkeit im Lesen, Schreiben, Buchstabieren sowie im Ausschneiden von Buchstaben und Ziffern aus Papier verfügte. Arnold weist darauf hin, daß vorzeitige Reife umschriebener Partialaktionen, welche den Eindruck einer überlegenen Intelligenz oder gar eines sich frühzeitig manifestierenden Genies erwecken kann, den Verdacht auf einen ersten pathologischen Prozeß psychischer, zerebraler oder endokriner Natur zuläßt.

Grewel und Margadant-Mainz analysieren einen Fall von *Lese- und Rechtschreibschwäche* und betonen, daß man sich mit der Feststellung einer «Entwicklungsdyslexie» nicht zufriedengeben darf. In jedem Fall wird eine gründliche neurologische und experimental-psychologische Untersuchung gefordert. Bei dem beschriebenen 9jährigen Knaben konnte auf diese Weise festgestellt werden, daß der «Leseschwäche» in diesem Fall eigentlich gar keine Lesestörung zugrunde lag, sondern eine ungenügende akustische Diskriminationsfähigkeit, Aufmerksamkeitsschwäche sowie ein unzureichender innerer Sprachaufbau. Diese Funktionen wurden beim Lesen und Schreiben so in Anspruch genommen, daß für das Verstehen des Gelesenen im Prozeß der distributiven Apperzeption keine Energie mehr übrigblieb.

Frau Borel-Maissonny beleuchtet aus *psychologischer Sicht Vorgänge bei Zuständen krankhaft gestörter Sprachentwicklung*. Dabei zeigt die Pathologie die Prozesse der Integration in verlangsamttem Ablauf, wodurch die normalen Mechanismen erhellt werden können. Eines der mitgeteilten Beispiele betrifft eine 20jährige Patientin mit zentralen Sprach-



entwicklungsstörungen im Gefolge einer beidseitigen hochgradigen Schwerhörigkeit. Ein otochirurgischer Eingriff verbesserte zwar das Hörvermögen beträchtlich, ohne jedoch die Fähigkeit zur Integration der Sprache zu verbessern. Man mußte auf verschiedene analytische Verfahren zurückgreifen, um ihr die Rede verständlich zu machen. Die Desintegrationen der Sprache gehorchen indes andern Gesetzen, und der Aufbau der Sprache zieht aus analytischen Rehabilitationsmethoden nur wenig Nutzen. Die Sprachfunktionen haben die Tendenz, sich global wiederherzustellen. Die vorliegende Arbeit ist von besonderem Interesse für den Phoniater, der sich mit Hörtraining beschäftigt, einem Gebiet, welches bei uns zurzeit im Vordergrund des Interesses steht. Die Ergebnisse decken sich mit unsern Erfahrungen bei der «Hörerziehung» nach hörverbessernden Operationen und unterstreichen den Grundsatz, daß die Verbesserung des Gehörs für den gestörten Spracherwerb alleine nicht ausreicht, sondern von einem psychologisch und pädagogisch ausgerichteten Hörtraining ergänzt werden muß.

*Van Riper* bespricht die *Faktoren, welche dem Stammer die Selbstkontrolle erschweren*. Er nimmt an, daß beim Kind die akustische Selbstkontrolle, beim Erwachsenen hingegen propriozeptive Rückkoppelung überwiege. Dementsprechend wird für die Sprachbehandlung vorgeschlagen, die eigene Lautung des Patienten und die des Therapeuten gleichzeitig zum Vergleich und zur Korrektur dem Patienten dichotisch zuzuleiten. Später sollte der neuerworbene Laut unter Vertäubung geübt werden, so daß die propriozeptive Kontrolle wiederhergestellt wird. Während ähnliche Verfahren bei der Behandlung von Stottern auch in Deutschland teilweise schon angewendet werden, sind sie nach Wissen des Referenten bei uns in die Stammerbehandlung noch kaum eingedrungen. Die Ausführungen von *Van Riper* sind geeignet, den Logopäden zur Auseinandersetzung mit der Gedankenwelt der Kybernetik anzuregen und stellen eine Bereicherung unserer therapeutischen Möglichkeiten bei der Behandlung von Stammer dar, insbesondere in solchen Fällen, in denen eine akustische Agnosie oder ähnliche Zustände bei der Entstehung des Stammelns im Spiele ist.

Nach *Beebe* und *Froeschels* ist allein aus der Beachtung gewisser Stottersymptome eine differentialdiagnostische *Abgrenzung des genuinen Stotters vom Nachahmungsstottern* möglich. Nachahmungsstottern wird oft verkannt.

*Freund* lenkt die Aufmerksamkeit auf das interessante Phänomen des *Stotters im Traume* und erörtert die Gründe, warum die Psychoanalyse bisher zu diesem Gebiete so wenig beigetragen hat.

Der führende Sprachpsychologe *Friedrich Kainz* erläutert einige Möglichkeiten zur *Erleichterung der Sprechhandlung*, wie sie instinktiv angewendet werden, wenn der normalerweise glatt verlaufende Sprechvorgang aus irgendwelchen Gründen erschwert oder im Extremfall pathologisch gestört ist. Bei Störungen der Artikulomotorik, die meist durch zentrale Hemmungen mitverursacht sind, wird vorzugsweise das Hilfsmittel des Stakkatosprechens, des Wort- und Silbendivisionismus sowie der Pausensetzung zum Einsatz gebracht. Wenn hingegen durch zögernde Gedankenarbeit das innersprachliche Geschehen nicht rasch genug ein Diktionskonzept zur Verfügung stellt, kommt es häufig zu einer Dehnung der Rede sowie zum Ausfüllen der Leerstellen durch Embolophrasen und «Embolophonien».

*Kandler* möchte der angewandten Sprachwissenschaft einen gebührenden Raum in der Sprachheilkunde eingeräumt wissen und betont, daß der *Brückenschlag zwischen Linguistik und Phoniatrie* zwar oft berufen, jedoch vielleicht zu selten verwirklicht wurde. Er zeigt 3 Stufen der Zusammenarbeit auf (1. Mitteilung sprachwissenschaftlicher Ergebnisse an den Fachmediziner, 2. Mitarbeit des Sprachwissenschaftlers an spezifisch sprachlichen Aufgaben der Medizin, z. B. am Aphasieproblem, und 3. Mitwirkung des Sprachwissenschaftlers an der tieferen Deutung von Sprachstörungen). So konnte beispielsweise die linguistische Analyse einen Fall von angeblich gestörter Raumorientierung als Amnesie deutschsprachlicher Dimensionssysteme aufklären. Die angeführten Beispiele belegen die Notwendigkeit einer Zusammenarbeit und lassen erkennen, daß beide

Disziplinen, die Medizin und die Linguistik, durch Nutzung der gegenseitigen Möglichkeiten fruchtbare Ergebnisse erhoffen dürfen.

*Van Thal* nimmt zu Fragen der *Ausbildung von Logopäden* Stellung.

Zur *Beurteilung einer stimmlichen Leistung* reicht die phonetische und laryngologische Untersuchung nicht aus. Dies hebt *Tarneaud* hervor. Er fordert eine phoniatische Untersuchung, die das psychologische, musische und physiologische Moment in gleicher Weise berücksichtigt.

Ein Beitrag zur *Taubstummepädagogik* schließt das Werk ab. *Zaliouk* beschreibt *Stimmstörungen bei Taubstummten*, welche als Folge der fehlenden Rückkoppelung und damit der defekten Selbstkontrolle verstanden werden. Die Stimme wird dadurch hoch, schrill und kann sich bis zu einem spastischen Falsett steigern. Zur Behandlung der hartnäckigen Falsettstimme hat *Zaliouk* eine vornehmlich taktil orientierte Methode entwickelt, welche hier beschrieben wird.

Trotz der Fülle des Gebotenen und der Verschiedenartigkeit einzelner Beiträge tritt uns das vorliegende Werk als geschlossenes und abgerundetes Ganzes entgegen, da sich alle Aufsätze unter den eingangs charakterisierten drei Hauptaspekten einordnen. Der Leitgedanke, daß jede Arbeit die Querverbindung zu einer Nachbarwissenschaft herstellen soll, sichert dem Werk eine vielseitige Verankerung im Raum der jüngsten wissenschaftlichen Forschung. Man darf dem Band bestätigen, daß er einen Querschnitt durch den derzeitigen Stand der Probleme in der Phoniatrie und Logopädie vermittelt.

A. Schilling, Marburg a. d. Lahn

*Raoul Husson: La voix chantée* (Commande cérébrale des cordes vocales; Classification chronaximétrique des tessitures; Mécanismes protecteurs du larynx; Analyse physiologique des techniques vocales et des grandes méthodes pédagogiques). Avec une Préface de M. le Dr *André Moulonguet*, membre de l'Académie de Médecine. Herausgegeben von Gauthier-Villars, Paris 6<sup>e</sup> 1960. XVI + 205 p., 96 Abb., ausführliche Bibliographie, broschiert, 23,00 NF.

Die Schrift des bekannten Pariser Forschers auf dem Gebiete der Stimmgebung repräsentiert die Synthese der in den letzten zwanzig Jahren und besonders intensiv seit 1952 den Problemen der Stimme beim Sprechen sowie beim Gesang gewidmeten Arbeiten des Autors. Diese Arbeiten bringen zwar individuelle, manchmal von der tradierten klassischen Lehre über die Art der Tätigkeit des Sprechorgans abweichende Ansichten des Autors, stützen sich jedoch in jedem einzelnen Falle auf eine gewissenhaft durchgeführte experimentelle Untersuchung, größtenteils unter Mitarbeit von Fachleuten auf dem Gebiet der Physiologie. Die Grundthese des Autors bildet die Lehre, daß die «Schwingungen» der menschlichen Stimmbänder nicht durch den Druck subglottischer Luft erzwungene Bewegungen sind, sondern durch vom Gehirn geregelte neuro-muskuläre Aktivität hervorgerufen werden. *Hussons* Feststellungen bedeuten naturgemäß einen tiefen Eingriff nicht nur in die bisherige Lehre über die Tätigkeit der Stimmbänder, sondern auch in die Behandlung von Stimmdefekten.

Diese seine erste synthetisch eingestellte Schrift widmete *Husson* der Gesangstimme und richtete sein Augenmerk vor allem auf die Bedürfnisse des Sängers und des Gesangspädagogen. Er wollte auf diese Weise den zahlreichen Künstlern aus der Réunion der Théâtres Lyriques Nationaux français und den Schülern und den zahlreichen Gesangsprofessoren an dem Conservatoire de Paris, welche sich für seine Versuchsarbeiten zur Verfügung stellten, seinen Dank ausdrücken. Trotz dieser seiner Einstellung zu der Gesangs- und der pädagogischen Praxis erleidet seine Fachlichkeit jedoch keinen Abbruch, so daß sowohl die Physiologen als auch die Ärzte von der Schrift mit Erfolg Gebrauch machen können. Sie ist und bleibt vor allem eine Schrift über Physiologie und Akustik der Phonation.

Der erste Teil von *Hussons* Schrift enthält in 4 Kapiteln die wesentlichsten Erkenntnisse aus den neuesten Forschungsergebnissen über Phonation; er ist daher der umfangreichste (er nimmt beinahe 100 S. ein). Im ersten Kapitel wird das Funktionieren der Stimmbänder erläutert, deren periodisches Öffnen durch Nervenimpulse geregelt wird. Es werden hier die wichtigen Forschungsergebnisse von *Goertler*, *Moulouguet*, *Portmann*, *Krmpotić*, *Fabre*, *Sabouroud*, *Gremy* usw. erwähnt. Im zweiten Kapitel werden die Teilnahme und der Einfluß der supraglottischen Höhlen bei der Phonation erörtert, und zwar in Anknüpfung an die neuesten Forschungsergebnisse des französischen Physikers *Y. Rocard*. Das dritte Kapitel behandelt die inneren Empfindungen des Sängers beim Gesang, die bisher des öfteren als Anzeichen von «Psychopathien» angesehen wurden; diese Empfindungen ermöglichen dem Subjekt vorteilhafte tonische Reflexe und tragen zur Vergegenwärtigung des von *Soulairac* 1955 definierten sogenannten «schéma corporel vocal» bei. Das vierte Kapitel ist der Phonationsphysiologie des *Nervus recurrens*, des motorischen Nerven der Stimmbänder, gewidmet. Die Exzitabilität dieses Nerven bildet die Grundlage für die Klassifizierung des Sängers vom Gesichtspunkt der Stimmhöhe (während das Timbre und die Intensität in dieser Richtung unabhängig bleiben). Diese Konsequenz der Forschungsergebnisse *Hussons* wird manchmal als zu spekulativ angesehen; in Wirklichkeit kann sie direkt in der Praxis ausgenutzt werden (so hat in diesem Jahr das Moskauer Staatskonservatorium einen Chronaximeter für Zwecke der Gesangspädagogik angekauft).

In den angeführten Kapiteln sind neben anderen zwei «mécanismes protecteurs» des Stimmorganismus klargestellt. Der erste von diesen Organismen ist die Deckung offener Laute («converture des sons ouverts»), den Sängern unter der Bezeichnung «passage» bekannt. Der zweite ist «l'impédance ramenée» (die rückwirkende Impedanz) am Kehlkopf beim Gesang, unter deren Einwirkung sich die Bildung von Vokalen gemeinsam mit der Tonhöhe befindet.

Der zweite Teil bringt in sechs Kapiteln eine ausführliche und durchgearbeitete physiologische Analyse aller möglichen Arten der Vokaltechnik. In der Konzeption *Hussons* ist die Vokaltechnik in erster Linie durch die neuro-motorischen Bedingungen, ferner durch die akustischen Modifikationen, durch verschiedenartige Ausnutzung der Schutzmechanismen (durch die Deckung offener Laute und die rückwirkende Impedanz des supraglottischen Raumes auf den Larynx), durch verschiedene Abstufung der inneren Empfindungen und schließlich durch ein differenziertes körperliches Vokalschema als dem Regulator der Gesangsleistung des Sängers charakterisiert.

Einen wichtigen Faktor bildet die Art, wie das Subjekt den supraglottischen Raum ausnutzen kann; nach diesem Kriterium kann die Vokaltechnik in drei Arten eingeteilt werden: 1. die Technik mit einer mäßigen Impedanz (gewöhnlich ungeschulte Stimmen), 2. die Technik mit einer starken Impedanz (mächtige Theaterstimmen) und 3. die mit einer geringeren oder größeren Nasalisierung verbundene Technik (bis zu den Grenzen eines pathologischen Fungierens).

Eine außerordentliche Aufmerksamkeit wird in der Schrift dem Gesang von mächtiger Leistungsfähigkeit (dem Theatergesang) gewidmet. Die an die Bühnensänger gestellten Anforderungen sind hier unter gleichzeitiger Anführung der geeigneten Technik in dem durch die anatomisch-physiologischen und neuro-endokrinen Verhältnisse des Objektes gegebenen Ausmaße ausführlich analysiert. Dieses Kapitel sollte die Aufmerksamkeit aller derjenigen finden, welche ihre Stimme bei öffentlichen, namentlich künstlerischen Darbietungen benutzen, und nach der Meinung *Hussons* auch aller, die Opfer einer unrichtigen Gesangserziehung behandelnden Phoniater.

Der dritte Teil der Schrift *Hussons* bildet einen wertvollen Behelf für die Gesangspädagogen, indem er eine mit physiologischer Auswertung verbundene Beschreibung aller wichtigsten, bis heute bekannten gesangserzieherischen Methoden liefert. Mit Rücksicht auf ihre große Menge teilt sie der Autor in fünf Gruppen ein. In die erste reiht er die Methoden der direkten Beeinflussung der Phonationsmuskulatur (appoggio,



Armins «Stauprinzip» usw.) ein, welche sich durch die Einstellung auf die, sei es nun inspiratorische oder expiratorische Atmungsart, Hebung oder Senkung der Lage des Kehlkopfs, die Gestaltung des Mundraumes und die Artikulationsbewegungen realisieren. Zu der zweiten Gruppe zählt *Husson* jene Methoden, deren Wesen in der direkten Einwirkung auf das vokalische Timbre besteht; hier führt er an erster Stelle die ausgezeichnete Methode *Labriets* mit der sogenannten «Vokalkompensation» an. In der dritten Gruppe befinden sich Methoden, welche die klar zum Bewußtsein kommenden und gut lokalisierbaren subjektiven Empfindungen («sensibilités internes») ausnutzen; die vierte enthält eine Übersicht der Methoden, die auf die Willensexpressivität des Timbres zu dem Zwecke eingestellt sind, durch die Regulierung der Spannung im Kehlkopf und die Gestaltung des supraglottischen Raumes eine geeignete Modifikation des Timbres hervorzurufen. Und schließlich die letzte, fünfte Gruppe behandelt einen bisher pädagogisch nicht ausgenutzten Mechanismus, der sich der Willenstätigkeit des Subjekts entzieht; es handelt sich um die Rückwirkung, die hervorgerufen wird durch auditive Anregungen des Subjekts beim Singen (in neuester Zeit durch die Versuche von *Tomatis* 1954 erwiesen). Damit schließt die Schrift *Hussons*.

Im ganzen kann gesagt werden, daß das Buch *Hussons* einen großen Beitrag für die wissenschaftliche Forschung sowohl über die physiologische Grundlage als auch den künstlerischen Aufbau der Gesangkunst bedeutet. Gestützt auf verläßlich durchgeführte Experimentalmethoden, bringt es in gedrängter Form die Ergebnisse unermüdlicher und durchwegs ganz origineller Arbeiten sowohl des Autors selbst als auch einer zahlreichen Gruppe französischer Forscher. Es wird an den Lehrern der Gesangkunst sowie an den Sängern selbst liegen, ob sie es zuwege bringen, aus dem Buch die sowohl für die Pädagogik des künstlerischen Gesangs als auch für die Gesangsleistung selbst nötigen Erkenntnisse abzuleiten.

B. Hála, Prag

*Hussons* Abhandlung stützt sich auf eine Reihe von Experimenten, welche mit Hilfe einer modernen elektrotechnischen und elektronischen Apparatur, wie z.B. der Einrichtungen zum Studium des *Tomatis*-Phänomens, des Glottographs von *Fabre* usw. durchgeführt werden. Die Funktion der in Frage kommenden Organe wird sowohl vom physiologischen Gesichtspunkt auf der einen Seite als auch vom Gesichtspunkt der physikalischen Akustik auf der anderen Seite beschrieben, und zwar mit Hilfe der modernsten Begriffe, wie zum Beispiel der Applikation der akustischen Impedanz. Daraus ist ersichtlich, daß der Autor auch auf dem Gebiet der physikalischen Akustik auf der Höhe steht und daß er ihre neuesten Erkenntnisse sehr gut beherrscht.

J. B. Slavík, Prag

W. R. Lee: **An English Intonation Reader.** Macmillan & Co., London 1960. 123 S., geb., 8 s. 6 d.

Nach langjähriger Unterbrechung – *H. Palmers* Bücher wurden erstmals 1922, *Armstrong* und *Wards Handbook* 1926 veröffentlicht – sind in letzter Zeit eine ganze Anzahl Lehrbücher der englischen Intonation erschienen. Die einen verlegen das Schwergewicht auf die systematische Einübung der verschiedenen *tone-patterns*: *W. St. Allen*: *Living English Speech* (1954); *R. Kingdon*: *The Groundwork of English Intonation* (1958); *M. Schubiger*: *English Intonation* (1958). Andere sind vor allem *Readers*, d. h. sie enthalten zusammenhängende Lesestücke mit Intonationszeichen: *P. A. D. MacCarthy*: *English Conversation Reader* (1956); *R. Kingdon*: *English Intonation Practice* (1958). Auch der Anteil linguistischer, d. h. nicht direkt der Spracheinübung dienender Betrachtung ist von Fall zu Fall verschieden. Während z. B. *Allen* ein reines Übungsbuch vorlegt, unterzieht *Kingdon* die englische Intonation einer weitausholenden Betrachtung.

Wie schon aus dem Titel hervorgeht, verfolgt das hier angezeigte Werk in erster Linie praktische Ziele. In einem ersten, «*Reading the Tone-Marks*» betitelten Abschnitt



wird der Lernende in die Benutzung des Buches eingeführt. Anhand eines kurzen sowohl mit Interlinearaufzeichnung als auch mit in den Text eingesetzten Tonzeichen versehenen Gesprächs lernt er diese Zeichen richtig zu interpretieren. Die Aufzeichnung ist einfach und leicht verständlich. Die Vereinfachung gegenüber *Kingdon* besteht vor allem darin, daß bei den Gleitlauten sowohl die Unterscheidung hoch/tief als auch die zwischen emphatisch und unemphatisch wegfällt. *Lee* verwendet nur vier Gleitlautzeichen:  $\backslash$ ,  $\uparrow$ ,  $\vee$  und  $\wedge$ <sup>1</sup>. Ein paar weitere Zeichen:  $\uparrow$ ,  $-$  und  $_$  weisen auf die *static tones* hin, ferner ein hochgestelltes  $\uparrow$  auf einen Gleitlaut in schwachbetonter Endsilbe. – Anhand von langen Reihen gleich zu intonierender Wörter und Sätzchen muß sich der Schüler jetzt die Gleitlaute aneignen. Die *static tones* werden – wie bei *Kingdon* – gesondert eingeübt. An einer Stelle (S. 20), so will es der Rezensentin scheinen, wird hier dem Lernenden durch die graphische Darstellung eine unnötige Schwierigkeit zugemutet.  $\vee$ *Mother*,  $\vee$ *which one*,  $\vee$ *journey*,  $\vee$ *recognized* usw. haben höchst selten die Intonation  $\vee$  oder  $\vee$ , sondern meist die viel leichter zu erlernende Form  $\vee \uparrow$  oder  $\vee \uparrow$ <sup>2</sup>.

Nach Absolvierung dieses «Vorkurses» kann sich der bloß auf die Spracherlernung eingestellte Ausländer gleich den Lesestücken zuwenden. Diese bilden eine «*A Week-End in the Country*» betitelte Folge von zehn Gesprächen<sup>3</sup>. Wer sich darüber hinaus für die Leistung der Intonation als sprachliches Ausdrucks- und Verständigungsmittel interessiert, findet im Kapitel «*Intonation at Work*» mannigfache Belehrung. *Lee* unterscheidet zwischen subjektiver und objektiver Funktion der Intonation. Erstere besteht darin, daß die Stimmführung, zusammen mit anderen Ausdrucksmitteln, wie Gesichtsausdruck, Gebärden und Lautstärke, die Haltung des Sprechers zum Ausdruck bringt. Sie wird vorerst nur kurz gestreift, erhellt dann aber deutlich aus den Kapiteln, in denen die Intonation der verschiedenen Satzarten besprochen wird. Am ausführlichsten wird der Fragesatz behandelt (S. 33–49), vor allem die *tail-question*, deren *tone-patterns* und die damit verknüpften Bedeutungsnuancen sich leicht darstellen lassen. Auch bei *Allen* nimmt *the question-tag* einen breiten Raum ein. Auf die Aussage-, Befehls- und Ausrufsätze entfallen dagegen nur insgesamt vier Seiten. *Lee* hätte dem Lernenden auch hier etwas mehr unter die Arme greifen dürfen<sup>4</sup>.

Die objektive Leistung der Intonation ist nach *Lee* eine zweifache. I. Die in einem gewissen Zusammenhang wichtigen Wörter werden durch die Stimmführung hervorgehoben, z. B.  $\backslash$ *John is writing to his father*; *John is \writing...* usw. Dem Deutschen begegnet hier vieles, was ihm aus der Muttersprache vertraut ist. II. Aus der Stimmführung erhellt oft das Verhältnis der Satzteile zueinander und folglich die Bedeutung des Satzes. Nicht selten hat der gleiche Satz, je nach Intonation, zwei ganz verschiedene Bedeutungen. Hier macht *Lee* den Studierenden auf Unterscheidungen aufmerksam, die er schon wiederholt besprochen hat<sup>5</sup> und die auch schon von anderen behandelt worden sind. Es handelt sich um folgende Fälle: 1.  $\uparrow$ *I didn't visit the \(-\) doctor because I was \(\vee\) ill*. 2. *The \(\uparrow\) men who were \(\vee\) lazy \(\uparrow\) stayed in \(\backslash\) bed*. 3. *They'd be \(\uparrow\) miserable if they were \(\uparrow\) rich*. 4. *She \(\uparrow\) looked \(\backslash\) up \(\uparrow\) happily*. 5. *She \(\uparrow\) doesn't speak to \(\vee\) anybody*.

<sup>1</sup> Man fragt sich, warum die am Schluß steigende Entsprechung von  $\wedge$ , nämlich  $\wedge$ , nicht einbezogen wird;  $\wedge$  bietet kaum größere Schwierigkeiten als  $\wedge$ , zumal wenn sich die Tonbewegung über mehrere Silben erstreckt.

<sup>2</sup> Siehe *Kingdon*: Groundwork, S. 11.

<sup>3</sup> Zu den Lesestücken sowie zu einem Teil des Einführungskurses ist ein Satz von sechs Linguaphonplatten erhältlich. An den Gesprächen sind acht Personen, vier Damen und vier Herren, beteiligt.

<sup>4</sup> Die sehr häufigen, also auch in den Lesestücken immer wieder vorkommenden Aussagesätze mit am Schluß fallend-steigender Intonation werden z. B. gar nicht berührt. Nur bei den Antworten auf *question-tags* wird u. a. die Intonation  $\backslash$ *No, it isn't* angeführt.

<sup>5</sup> *Maitre phonétique*, 1953, 1955, 1956. *Lingua* V, 1956.

Bei allem Interesse, welches *Lees* theoretische Ausführungen zu wecken vermögen, ist dieser Teil des Buches doch wenig befriedigend. Er ist nicht klar genug aufgebaut. Zusammengehöriges wird oft getrennt, Verschiedenartiges vermengt. Hier ein paar Beispiele: Der oben unter (2) angeführte Unterschied zwischen einschränkenden und erweiternden Relativsätzen wird richtig als objektive Leistung der Intonation gewertet. Doch die entsprechende Unterscheidung bei der Apposition ( $\uparrow(\backslash)$  *Smith the bank manager*) steht im Kapitel, in dem die Intonation der Satzarten zur Sprache kommt. Unter dem Titel «*Co-ordination*» erscheinen nicht nur Sätze wie:  $\uparrow$  *Nobody came and \nothing happened*, sondern erstaunlicherweise auch *We've been waiting for you at the station*. Unter «*Reservation*» finden wir neben *You can see it occasionally* den Fragesatz *Where were you on Friday?*, den man nicht «a question... ending in a fall qualified by an addition» (69) nennen kann. – Solche Begriffsverwirrungen sind um so bedauerlicher, als das Buch als Ganzes gegenüber anderen *Intonation Readers* einen großen Vorteil hat: Die theoretische Betrachtung und die Lesestücke sind aufeinander abgestimmt. Geschickt sind die Verwendungsmöglichkeiten der *tone patterns* in den verschiedenen Satzarten und auch die gelegentlich damit verbundenen Bedeutungsunterschiede in den Text eingeflochten. Und trotzdem wirken die Gespräche nicht gekünstelt.

Eine Einzelbemerkung: Seite 31, Zeile 2, sollte es heißen: “A stress difference helps to distinguish *in'sult* (verb) from *'insult* (noun) ... No matter how we say these words, the distinction is clear if we keep to the *tonetic* stress-pattern (nicht bloß *stress pattern*).” Erwiesenermaßen kommt dem dynamischen Akzent nur im Zusammenhang mit der entsprechenden Intonation differenzierende Bedeutung zu. – Druckfehler: S. 110 (Mitte): *pesimizm*, nicht *pezimizm*; S. 113 (Zeile 6): əɪʃɔ:, nicht əʃɔ:; S. 118 (fünfte und vierte Zeile von unten): zwei Schlußpunkte fehlen. Maria Schubiger, Basel

*Milan Romportl: Zvuková stránka souvislé řeči v nářečích na Těšínsku* («Die klangliche Seite der zusammenhängenden Rede in den Dialekten des Teschener Gebiets»). Publikace Slezského ústavu ČSAV 23. Krajské nakladatelství v Ostravě, Ostrava (Ostrau) 1958. 125 S. und 2 Karten.

Im Jahre 1940 wurde eine nachgelassene Studie von *Stanislav Petřík* über die Satzintonation in den mährisch-schlesischen Mundarten veröffentlicht<sup>1</sup>, die nach dem Willen des Verfassers durch Schilderungen verschiedener Eigentümlichkeiten und auffälliger Erscheinungen sowie durch Hinweise auf die Vielfältigkeit der Problematik einen Ausgangspunkt für weitere Forschungen bilden sollte. Es ist *Romportls* Verdienst, die Anregungen *Petříks* aufgegriffen und eine der Dialektgruppen, nämlich die schlesisch-polnischen Mundarten in dem tschechoslowakischen Teil des Teschener Gebiets (Schlesien), einer systematischen satzphonetischen Durchforschung unterzogen zu haben. Nach zwei vorbereitenden Abhandlungen, die neben einer Übersicht der Akzentverhältnisse und Anmerkungen zur phonetischen Realisierung der phonologisch irrelevanten Vokalquantität vor allem eine Darstellung der satzmelodischen Gegebenheiten in verschiedenen Gebieten der erwähnten Dialektgruppe anstreben<sup>2</sup>, legt der

<sup>1</sup> *St. Petřík: Zur Satzintonation der mährisch-schlesischen Mundarten. Slavia 17: 481–548 (1939/40).* – Es ist darauf hinzuweisen, daß der Terminus «Satzintonation» bei *Petřík* die Gesamtheit der Akzent-, Tonverlaufs- und Dauerverhältnisse auf der Satzebene umfaßt.

<sup>2</sup> *M. Romportl: Přízvuk, kvantita a melodie v nářečí na Jablunkovsku* («Akzent, Quantität und Melodie im Dialekt des Jablunkauer Gebiets»). *Slezský sborník* (Acta silesiaca) 52: Beilage des Heftes 1/2, 48 S. (1954). – *Přízvuk a melodie nářečí na Těšínsku* («Akzent und Melodie der Dialekte des Teschener Gebiets»). *Adolfu Kellnerovi sborník jazykovědných studií*, pp. 53–78 (Slezský studijní ústav, Opava 1954).

Prager Dozent mit der oben angeführten Arbeit nunmehr die Summa seiner Untersuchungen vor.

Die vorliegende Arbeit stellt in mehrfacher Hinsicht einen willkommenen Beitrag dar. Im Bereich der slavischen Sprachen und ihrer Mundarten (durchaus aber nicht in diesem Bereich allein) besteht an satzphonetischen Untersuchungen noch ein spürbarer Mangel. Zudem befindet sich die phonetische Erforschung der zusammenhängenden Rede (die Satzphonetik) selbst noch in den Anfangsstadien ihrer Entwicklung, und in dieser Phase ist jede Arbeit für den weiteren Fortschritt wesentlich, sei es dadurch, daß sich vielleicht manche Prolegomena zu einer späteren einheitlichen Methodik finden lassen, sei es durch die generellen Aspekte der Resultate. Auch wenn man sich weder speziell mit den sogenannten tschechisch-polnischen Übergangsmundarten noch allgemein mit dem Westslavischen oder vielleicht mit dem Slavischen überhaupt befaßt, wird man durch die Arbeit *Romportls* zu manchen Überlegungen und Gedankengängen angeregt werden.

Die Struktur der Arbeit ist schnell skizziert. Nach einem einleitenden Kapitel, in dem der Verfasser den Gegenstand seiner Untersuchungen und seine Arbeitsweise erläutert (S. 5–13), folgt je ein Kapitel über Akzent (S. 14–50), Quantität (S. 51–63), Satzmelodie (S. 63–79) sowie über Sprechtempo und Redepausen (S. 80–86). In einem weiteren zusammenfassenden Abschnitt versucht der Autor sodann eine Synthese zu geben (S. 87–92). Dialekttexte aus dem untersuchten Gebiet (die äußerst sorgfältig durchgeführte phonetische Transkription verdient hervorgehoben zu werden), ein russisches und ein deutsches Résumé sowie ein gemischtes Sach- und Namenregister schließen die Arbeit ab.

Gemäß der von der Prager phonetischen Schule angewandten Terminologie<sup>3</sup> nennt *Romportl* als Gegenstand seiner Untersuchungen an den schlesisch-polnischen Mundarten die *Modulationen der zusammenhängenden Rede* oder *Sprechmodulationen* und versteht darunter solche lautlichen Erscheinungen, die den höheren Einheiten des lautlichen Baues der Sprache, dem Takt (d. h. Akzenttakt) und dem Satzabschnitt bzw. dem Satz zugeordnet sind, mit diesen Einheiten eng zusammenhängen und sie bestimmen und abgrenzen oder zumindest abzugrenzen und zu bestimmen helfen (S. 5)<sup>4</sup>. Dieser Definition mögen gleich einige Bemerkungen folgen.

Wenn der Verfasser Silbe, Akzenttakt und Satzabschnitt bzw. Satz «höhere Einheiten des lautlichen Baues» nennt und an anderer Stelle die Laute als Grundelemente des lautlichen Baues der Sprache bezeichnet (S. 6), so erinnert dies an das ältere Modell der Phonetik, in dem die Laute als Elemente der Rede begriffen und alle höheren Einheiten, wie Silbe, Akzenttakt oder Akzentgruppe und Ausspruch bzw. Satz, aus ihnen abgeleitet wurden. Gerade die Satzphonetik aber sollte sich ganzheitlich orientieren! Ebenso wie die Linguistik dazu übergegangen ist, die Sprache nicht mehr von den kleinsten Elementen her aufzubauen, sondern den Satz als Sinn Ganzes zum Ausgangspunkt und zur Grundlage zu nehmen, sollte auch die Phonetik, ganz besonders natürlich bei Untersuchungen der zusammenhängenden Rede, bewußt vom Satz (hier im weitesten Sinne des Wortes gefaßt) als der eigentlichen Redeeinheit ausgehen. Denn alles Sprechen vollzieht sich – um eine Binsenwahrheit zu sagen – in Sätzen.

Wenn die «Sprechmodulationen» als lautliche Erscheinungen bezeichnet werden, die den höheren Einheiten des lautlichen Baues der Sprache zugeordnet sind, so trifft das zweifellos zu. Doch sollte man darüber nicht vergessen, *warum* dem so ist. Die moderne Sprachpsychologie lehrt, daß die mit den Wörtern als Zeichen verbundenen Bedeutungen im Satz in einer bestimmten Weise kombiniert, in einen Sinnzusammenhang

<sup>3</sup> Vgl. B. Hála: *Úvod do fonetiky* («Einführung in die Phonetik») (Melantrich, Prag 1948).

<sup>4</sup> Im deutschen Résumé (S. 110) spricht *Romportl* von «höheren Einheiten des phonetischen Systems der Sprache». (Kursivdruck im Zitat von mir – H.-W. W.)



eingebaut und zu einer Ganzheitlichkeit integriert werden<sup>5</sup>. Dieser Zusammenschluß zu einer sinnvollen Redeeinheit findet seinen lautlichen Ausdruck in den «Sprechmodulationen». Mit dieser Feststellung nähern wir uns dem funktionellen Aspekt, der beim Verfasser in der Formulierung anklingt, daß die den höheren Einheiten des lautlichen Baues der Sprache zugeordneten lautlichen Erscheinungen mit diesen Einheiten eng zusammenhängen und sie bestimmen und abgrenzen oder zumindest abzugrenzen und zu bestimmen helfen (s. o.). Damit können wir uns jedoch nicht zufriedengeben. Die «Sprechmodulationen» dienen der lautlichen Zusammenfassung des Satzkörpers (d. h. der einzelnen Wörter als Bedeutungsträger) zu einer gewissen geschlossenen Einheit (zur sinnvollen Redeeinheit). Als *Grundfunktion* der «Sprechmodulationen» ist zweifellos die *Umwandlung der Wörter als appellerativer Einheiten in die kommunikative Redeeinheit «Satz»* anzusehen<sup>6</sup>. Diese Funktion könnte auch *satzkonstituierende Funktion* genannt werden. Nun wird die Redeeinheit «Satz» mit Hilfe der «Sprechmodulationen» nicht nur konstituiert, sondern die Redeeinheiten werden im Zusammenhang der Rede voneinander abgegrenzt, und außerdem werden verschiedene Arten der Redeeinheiten voneinander unterschieden. Innerhalb der Grundfunktion wären also zwei *Teilfunktionen* zu unterscheiden: eine *delimitative* und eine *differenzierende*. Diese Teilfunktionen dürfte der Verfasser im Sinn gehabt haben, als er von «Abgrenzung» und «Bestimmung» sprach.

Hinweise auf die Umwandlung von Wörtern in sinnvolle und verständliche Redeeinheiten als die eigentliche Funktion der satzphonetischen Faktoren sind bereits bei dem in manchen seiner Gedankengänge sehr modern anmutenden *Eduard Sievers* zu finden. Es sei ein Zitat angeführt: «Unter einem Satz wollen wir hier eine jede selbständige gesprochene Äußerung verstehen, d. h. eine jede in sich geschlossene Lautmasse, die in einem bestimmten Zusammenhang, sei es der Rede, sei es der Situation überhaupt, einen bestimmten Sinn (Gedanken oder Stimmung) zum Ausdruck bringen soll und in diesem bestimmten Sinn von dem Hörer verstanden wird<sup>7</sup>.» Hervorzuheben ist die Formulierung «in sich geschlossene Lautmasse». In § 613 der «Grundzüge» (5. Aufl.) wird diese Formulierung durch die Feststellung näher erläutert, daß ein verständlicher, eindeutiger Satz erst dadurch entstehe, daß die Wörter in einer ganz bestimmten Weise zu einer phonetischen Einheit zusammengeschlossen werden. In § 617 werden sodann die Abstufung nach Stärke, Tonhöhe und Dauer als *Variationsmittel* genannt, die für die phonetische Charakteristik des Satzes (für die Zusammenfassung der Wörter zu einer «geschlossenen Lautmasse») in Betracht kommen. Erst der Zusammenschluß zu einer phonetischen Einheit mittels Abstufungen nach Stärke, Dauer und Tonhöhe – so ungefähr könnte man die *Sieversschen* Ausführungen kurz zusammenfassen – ver-

<sup>5</sup> Vgl. z. B. *Fr. Kainz*: *Psychologie der Sprache*; Bd. 1: *Grundlagen einer allgemeinen Sprachpsychologie* (F. Enke, Stuttgart 1941).

<sup>6</sup> In dem Aufsatz «Sentence intonation from a functional point of view», *Word* 16: 34–54 (1960), sagt *F. Daneš*: «The fundamental function of intonation is to transform words, as appellative units, into communicative units, i.e. into utterances. Each word or succession of words automatically becomes an utterance when it is pronounced in a certain intonational form» (op. cit. S. 43–44). – *Daneš* vereinfacht die Verhältnisse insofern, als er allein der Intonation die Rolle des transformierenden Faktors zuerkennt. An der Konstituierung der kommunikativen Redeeinheit «Satz» sind jedoch *alle* satzphonetischen Elemente (außer dem melodischen Element noch das dynamische und die quantitativen Elemente) maßgeblich beteiligt. Wenn auch dem melodischen Element in der Hierarchie der erste Platz gebühren dürfte, erst dem *Zusammenwirken* der genannten Elemente kommt die Funktion zu. (Mit der Arbeit von *Daneš* werden wir uns an anderer Stelle noch ausführlich beschäftigen.)

<sup>7</sup> Vgl. *E. Sievers*: *Grundzüge der Phonetik*; 5. Aufl. (Breitkopf & Härtel, Leipzig 1901).



wandelt ein Wort oder eine Reihe von Wörtern in einen verständlichen, eindeutigen Satz mit bestimmtem Inhalt. Die zentrale Bedeutung der «Sprechmodulationen» für die Konstituierung der Redeeinheit «Satz» war von *Sievers* im wesentlichen richtig erkannt worden.

Wenden wir uns nach diesem Exkurs wieder der vorliegenden Arbeit zu! Die Materialgrundlage der Untersuchungen dürfte als beispielhaft zu bezeichnen sein. Für die Arbeit wurde Material von insgesamt 104 Informanten männlichen und weiblichen Geschlechts aus 61 Ortschaften des untersuchten Gebiets ausgewertet. (Vgl. die Kartenskizze S. 8 und das Verzeichnis der Sprecher S. 10–11.) Das Material wurde vom Verfasser auf drei Reisen in die Gebiete der untersuchten Dialektgruppe (im Juni 1952, im September 1953 und im September 1955) gesammelt. Bei der Geländearbeit konnte ein Magnetophongerät eingesetzt werden. Die Vorteile einer solchen «Speicherung» des Materials sind hinreichend bekannt. Die Tonbandaufnahmen können beliebig oft abgehört und die auf dem Wege der auditiven Analyse gewonnenen Belege wiederholt auf ihre Genauigkeit hin kontrolliert werden. Zusätzlich eröffnet sich die Möglichkeit instrumenteller Analysen, die vom Verfasser auch genutzt wurde.

Mit 36 Seiten bildet die Darstellung der *Akzentverhältnisse* in den untersuchten Mundarten den umfangreichsten Teil der Arbeit. Hier ist besonders die sorgfältige statistische Arbeitsweise des Verfassers hervorzuheben, die nicht wenig zur Klärung verschiedener dringender Probleme beigetragen hat. *Petrík* korrigierte seinerzeit die herrschende Auffassung vom Vorletztzilbenakzent der schlesisch-polnischen Mundarten des Teschener Gebiets dahingehend, daß in drei- und mehrsilbigen Wörtern bzw. Akzenteinheiten die Pänultimabetonung (der «polnische» Typus) und die Anfangssilbenbetonung (der «tschechische» Typus) einander ablösen. Da er in dem von ihm konstatierten starken Schwanken der Akzentuierung die Tendenz erkannt haben wollte, daß nichthervorgehobene Wörter größtenteils eine schwache Anfangssilbenbetonung tragen, während wenigstens gering hervorgehobene Wörter die Vorletztzilbenbetonung und stark hervorgehobene wieder eine (diesmal jedoch starke) Anfangssilbenbetonung aufweisen, glaubte er von bestimmten Funktionen der verschiedenen Akzentuierungsweisen sprechen zu können. *Romportl* mußte sich also gewissermaßen zwangsläufig mit der Frage nach dem Verhältnis von Vorletztzilben- und Anfangssilbenbetonung und nach ihren möglichen sprachlichen Funktionen befassen. Es ergab sich, daß *Petrík* aus einem vielfach stark beschränkten Material etwas weittragende Schlüsse gezogen hatte, die sich anhand der Akzentstatistiken als unhaltbar erwiesen. Die Pänultimabetonung wird vom Verfasser nach wie vor als Grundtypus der Akzentuierung der untersuchten Mundarten angesehen. Die mit stärkerer oder schwächerer Frequenz auftretende Anfangssilbenbetonung stellt den Ausführungen *Romportls* zufolge eine Akzentverschiebung dar, deren Ursache in den dynamischen Verhältnissen des Satzes gesucht werden mußte. Vor allem sind natürlich Akzentverschiebungen, die im Zusammenhang mit einer Verstärkung des Akzents (mit dem sogenannten Hervorhebungsakzent) auftreten, als satzdynamische Erscheinung zu bewerten. Diese Kategorie, die allgemein in Sprachen bzw. Mundarten zu finden ist, deren Akzent keine phonologisch relevante (distinktive) Funktion zukommt, teilt Verf. von den anderen ab. Aber auch die interessanteren Akzentverschiebungen bei normaler (nicht verstärkter) Akzentuierung beruhen *Romportl* zufolge mehr auf der Satzodynamik als auf der Dynamik des Wortes.

Ferner konnte Verf. auf Grund seiner Akzentstatistiken nachweisen, daß Dreisilbler in einem weitaus stärkeren Maße zur Akzentverschiebung inklinieren, als es bei viersilbigen und längeren Wörtern oder Akzenteinheiten der Fall ist. Außerdem konnte überzeugend dargelegt werden, daß der Anfangssilbenakzent im südlichen Teil des untersuchten Gebiets eine weit höhere Frequenz als in den nördlicher gelegenen Teilen aufweist; und auch hier sind wiederum Dreisilbler für die Akzentverschiebung am anfälligsten. (Über die geographische Situation orientieren zwei Karten, die nach S. 40

bzw. 42 der Arbeit eingehftet sind.) In dieser Hinsicht hat *Romportl* Vorzügliches geleistet.

Im Hinblick auf das Vorhandensein der Anfangssilbenbetonung glaubt Verf. von einer Abschwächung der delimitativen Funktion des Pänultimaakzents sprechen zu können. Diese Abschwächung sei aber nur generell aufzufassen: das von Teilgebiet zu Teilgebiet des untersuchten Sprachraumes variierende Schwanken zwischen Pänultima- und Anfangssilbenbetonung sei nicht etwa gleichbedeutend mit verschiedenen Stufen der Abschwächung. Diesen Ausführungen vermögen wir nicht ganz beizustimmen. Wenn ein «delimitativer» Pänultimaakzent Worte oder Akzenteinheiten dadurch voneinander abzugrenzen vermag, daß er auf die Tatsache hinweist, daß das Ende des Wortes oder der Akzenteinheit nach einer weiteren, unbetonten Silbe folgt, dann grenzt der Anfangssilbenakzent ebenso gut dadurch ab, daß er auf der ersten Silbe eines Wortes bzw. einer Akzenteinheit ruht. Delimitiert wird also immer, wenn auch die Art der Abgrenzung verschieden ist. Man sollte infolgedessen wohl besser von einer unterschiedlichen funktionellen Belastung beider Akzentuationstypen sprechen. Gehen wir aber einen Schritt weiter! Ist denn die Funktion des Akzents, besser gesagt: die *primäre* Funktion, in den untersuchten Mundarten überhaupt «delimitativ»? Läßt die Tatsache, daß der Akzent in der zusammenhängenden Rede aus dynamisch-rhythmischen Gründen von der vorletzten auf die erste Silbe verschoben werden kann, ohne daß die andere Art der «Abgrenzung» zu irgendwelchen Störungen in der Verständigung führt oder zumindest innerhalb eines sprachlichen Kollektivs auffallend wirkt, nicht eventuell auf eine andere Funktion, nämlich die der dynamisch-rhythmischen Gliederung der Rede, schließen? Diese wie es uns scheint naheliegende Frage wurde vom Verf. nicht gestellt, obwohl doch gerade der Umstand, daß Dreisilbler für die Akzentverschiebung viel anfälliger sind als Vier- und Fünfsilbler, in denen dafür mit großer Regelmäßigkeit ein sogenannter Nebenakzent auf der ersten Silbe erscheint, noch besonders darauf hinweisen dürfte. Auch hätte eine Stellungnahme zu den Arbeiten *Galtons*<sup>8</sup> vielleicht manches zu einer Klärung dieser Problematik beitragen können.

Zur Frage des sogenannten «Satzakzents», der generell als eine Verstärkung der Akzentsilbe des sinnwichtigsten Wortes eines Satzes (des «psychologischen Prädikats» oder «Aussagekerns») aufgefaßt wird, konnte Verf. durch auditive Analyse mit Hilfe des Bandschnittverfahrens feststellen, daß normalerweise (wenn nicht bereits «Hervorhebungsakzent» vorliegt) die als «satzakzentuiert» wahrgenommene Silbe nicht nur nicht stärker, sondern sehr oft sogar schwächer als die anderen Akzentsilben des Satzes ist. Diese Situation entspricht durchaus den Verhältnissen in der tschechischen Schriftsprache<sup>9</sup>. Es scheint, daß der «Satzakzent» durch das Zusammentreffen mehrerer wichtiger Erscheinungen signalisiert wird. Als solche wären zu nennen: Stellung des «psychologischen Prädikats» am Satzende; Dehnung des Vokals in der Endsilbe des Satzes, d. h. in der Silbe, die unmittelbar auf die «satzakzentuierte» folgt; Verbindung der satzmelodischen Kadenz, d. h. des funktionell relevanten Abschnitts der melodischen Linie des Satzes, mit dem «psychologischen Prädikat».

Im Hinblick auf die durchschnittliche Stärke der Akzentuierung und auf das Stärkeverhältnis zwischen «akzentuierten» und «nichtakzentuierten» Silben werden die untersuchten Dialekte vom Verf. zwischen das Tschechische und das Russische eingestuft.

Es ist sehr zu begrüßen, daß der traditionell «dynamisch» genannte Akzent vom Verf. als Komplexqualität aufgefaßt wird. Obwohl sich *Romportl* nicht näher mit der

<sup>8</sup> *H. Galton*: On the supposed delimitative accent in West Slav. *Archivum Linguisticum* 7: 123–139 (1955). – Über das Prinzip rhythmischer Alternation im Slavischen. *Wiener slavist. Jahrb.* 5: 37–58 (1956).

<sup>9</sup> *A. Skaličková*: K otázce větého přízvuku v češtině («Zur Frage des Satzakzents im Tschechischen»). *Universitas Carolina, Philologica II, 1*: 55–66 (1956).

phonetischen Natur des Phänomens «dynamischer Akzent» befaßt (dieses Problem lag ja außerhalb des Rahmens seiner Arbeit), weist er doch ausdrücklich darauf hin, daß er unter dem Begriff der Stärke des Akzents nicht nur die physikalische Intensität versteht, sondern immer einen ganzen Komplex klanglicher Erscheinungen (S. 19).

Das Merkmal der vokalischen *Quantität* wird in den schlesisch-polnischen Mundarten des Teschener Gebiets, ähnlich wie im Polnischen und im Gegensatz zur tschechischen Schriftsprache und der Mehrheit der tschechischen Dialekte, nicht zu distinktiven Zwecken genutzt und ist infolgedessen phonologisch irrelevant. Um so eingehender ist natürlich zu prüfen, ob und in welchem Maße Veränderungen der Vokaldauer auf der Satzebene funktionell verwendet werden. Dadurch, daß Dauerunterschiede nicht als Realisierung einer bestimmten, der Wortunterscheidung dienenden sprachlichen Norm auftreten, läßt die phonologische Struktur der untersuchten Mundart grundsätzlich die Möglichkeit einer anderen Verwendung offen.

Verf. beschränkt sich nicht auf den funktionellen Aspekt, sondern gibt auch eine generelle Darstellung der Quantitätsverhältnisse. Seine Ausführungen basieren auf der Ausmessung elektromagnetischer kymographischer Registrierungen, die von Teilen der Magnetophonaufnahmen gewonnen wurden. Wir erfahren zunächst, daß die spezifischen Dauerverhältnisse der Vokalkategorien in den untersuchten Mundarten sich nicht von denjenigen der tschechischen Schriftsprache unterscheiden. Ein geschlossener Vokal ist unter im übrigen gleichen Bedingungen kürzer als ein offener. Die *Quantität* steht ferner in einem direkten Verhältnis zum Akzent. Die Vokale akzentuierter Silben weisen eine relativ längere Dauer als die der nichtakzentuierten Silben auf. Besonders deutlich ist diese Erscheinung bei einer Akzentverstärkung («Hervorhebungsakzent») zu beobachten. Die relativ geringste Dauer fällt gewöhnlich dem Vokal der unmittelbar auf die Hauptakzentsilbe eines Wortes (bzw. einer Akzenteinheit) folgenden nichtakzentuierten Silbe zu. Jeder Vokal ist außerdem um so kürzer, je länger die Einheit ist, der er angehört. Obwohl es sich bei den genannten Erscheinungen um Tendenzen allgemeinsprachlichen Charakters handelt, so daß experimentelle Untersuchungen bei gleichen sprachlichen Verhältnissen zu relativ gleichen Resultaten führen<sup>10</sup>, sind *Romports* Darstellungen als Ergänzung der bisherigen Literatur willkommen, liefern sie doch Quellenmaterial aus einem in dieser Hinsicht noch nicht näher untersuchten Sprachgebiet.

Wir kommen nun zu den (von der Satzebene aus gesehen) funktionellen Aspekten der Quantitätserscheinungen. (1) Der Vokal in der Endsilbe eines Satzes oder Satzabschnitts erfährt eine deutliche Längung. In diesen Fällen ist die Modifizierung der Vokalquantität an der Abgrenzung des Satzes als Redeeinheit beteiligt und übt im Zusammenhang mit anderen Faktoren (Satzmelodie und Redepause) eine delimitative Funktion aus. Das ist z. B. (um Beispiele aus dem Westslavischen zu nennen) auch im Polnischen<sup>11</sup> und Tschechischen<sup>12</sup> der Fall. (2) Eine starke Dehnung der Akzentsilben (vor allem der Akzentsilbe des sinnwichtigsten Wortes) pflegt auch bei emphatischer oder emotioneller Aussprache zu erscheinen. Leider werden diese Fälle vom Verf. zusammen mit der Längung infolge einer Akzentverstärkung bei einfacher (d. h. nicht-emotioneller Hervorhebung) behandelt. (Daß sogenannte «emotionelle» Dehnungen nicht von einer gleichzeitigen Akzentverstärkung begleitet sein müssen, beweist eins der vom Verf. angeführten Beispiele.) Die im Vergleich mit den anderen Abschnitten des Quantitätskapitels (besonders mit den Darstellungen der allgemeinen Tendenzen) sehr

<sup>10</sup> Vgl. B. Malmberg: Die Quantität als phonetisch-phonologischer Begriff (C. W. K. Gleerup, Lund 1944). Besonders S. 6–13.

<sup>11</sup> Zur Dehnung der «satzauslautenden» Silben im Polnischen vgl. J. Rozwadowski: Szkiz wymowy (fonetyki) polskiej («Abriß der Aussprache [Phonetik] des Polnischen»). Mat. i Prace Kom. Język. 1: 95 ff. (Krakau 1904).

<sup>12</sup> Vgl. Fußnote 9.



kurz gehaltenen Ausführungen sind zudem mit nur 5 Beispielen aus beiden behandelten Bereichen wirklich spärlich illustriert. Wir finden keinerlei Hinweis darauf, ob die «emotionellen» Dehnungen nicht im expressiven Bereich der Satzebene eine wesentliche Rolle spielen, ob sie nicht ein satzphonetisches Element darstellen, das zwar keinerlei Einfluß auf die Bedeutung des Satzes, wohl aber auf die *Bedeutsamkeit* hat. (In den westlachsichen Mundarten des Tschechischen, die ebenso wie die vom Verf. untersuchten schlesisch-polnischen Mundarten des Teschener Gebiets keine phonologisch relevante Quantität kennen und eine Pänultimabetonung aufweisen, besteht z. B. folgende Möglichkeit: Wird in einer Aussage durch eine spezielle melodische Form Erstaunen angedeutet, so kann durch «emotionelle» Quantität zusätzlich ein großes Maß des Erstaunens signalisiert werden. So kann eine relativ feststehende Satzbedeutung in ihrer Bedeutsamkeit variiert werden<sup>13</sup>.) Emphatische Veränderungen der Vokalquantität dürften nicht grundsätzlich als bloße Begleiterscheinungen einer emotionalen Aussprache gewertet werden. In *Romportls* Abschnitt «Dehnung infolge Hervorhebungsakzent und Emphase» (S. 56–58) ist der funktionelle Aspekt sichtlich zu kurz gekommen.

Daß der funktionelle Aspekt nur teilweise berücksichtigt wurde, zeigt auch die Gesamtanlage des Kapitels «Quantität», das folgende Gliederung aufweist<sup>14</sup>: A. «Methode der Untersuchung»; B. «Quantität in Wort und Satz», 1. «Verhältnis der Quantität verschiedener Vokale», 2. «Akzentuierte und nichtakzentuierte Vokale», 3. «Dehnung infolge Hervorhebungsakzent und Emphase», 4. «Quantität der Vokale in den Endsilben von Sätzen und Satzabschnitten», 5. «Einfluß der Länge des Taktes auf die Quantität der Vokale»; C. «Zusammenfassung». In Punkt 1, 2 und 5 des Abschnitts B werden allgemeinsprachliche Erscheinungen behandelt, Punkt 4 befaßt sich mit funktionell relevanten Gegebenheiten, in Punkt 3 wird beides vermischt. Eine vom funktionellen Standpunkt ausgehende Gliederung dürfte sich als vorteilhafter erwiesen haben. Die geringe Berücksichtigung des funktionellen Aspekts ist um so bedauerlicher, als *Romportl* selbst an anderer Stelle die Forderung erhoben hat: «Es wird also nötig sein, vorerst eine ausführliche und präzise Analyse der satzphonetischen Elemente und Mittel..., u. zw. wie vom phonetischen, so auch vom linguistischen (phonologischen) Standpunkt durchzuführen<sup>15</sup>.»

Das Material zur *Satzmelodik* der untersuchten Mundarten wurde vom Verf. bereits in den vorbereitenden Studien (vgl. Fußnote 2) weitgehend beschrieben und kommentiert. So gibt *Romportl* in dem Kapitel «Melodie» der vorliegenden Arbeit (S. 63–79) eine Synthese und versucht das satzmelodische System zu beschreiben.

Der Tonverlauf (die Melodie) eines Satzes übt dem Verf. zufolge in den untersuchten Mundarten einmal eine satzdifferenzierende Funktion aus, zum anderen werden mit seiner Hilfe noch die gefühlsmäßigen Stellungnahmen des Sprechers zu dem Gesagten ausgedrückt<sup>16</sup>. *Romportl* geht davon aus, daß diese Leistungen der Satzmelodie sich in drei Ebenen vollziehen. Durch unterschiedliche melodische Gestaltung wird ein nichtabgeschlossener (weiterweisender) Satz von einem abgeschlossenen differenziert (= 1. Ebene), ferner wird eine Entscheidungsfrage von einer Nichtfrage unterschieden (= 2. Ebene), und es wird schließlich ausgedrückt, ob es sich um einen gefühlsgefärbten oder emotional neutralen Ausspruch handelt (= 3. Ebene)<sup>17</sup>.

<sup>13</sup> Darüber ausführlich in einer Arbeit «Satzphonetik des Westlachsichen», die sich zurzeit im Druck befindet.

<sup>14</sup> Die Titel der Abschnitte wurden von mir übersetzt – H.-W. W.

<sup>15</sup> M. *Romportl*: Zum vergleichenden Studium der Satzphonetik. Z. f. Phonetik 10: 351–358 (1957). Das Zitat S. 352.

<sup>16</sup> Eingangs wurde bereits darauf hingewiesen, daß Verf. den «Sprechmodulationen» nur eine (wie er sagt) bestimmende und abgrenzende, keinesfalls aber eine satzkonstituierende Funktion zuerkennt.

<sup>17</sup> Zum Drei-Ebenen-Schema *Romportls* und seiner Entstehung vgl. H.-W. Wodarz: Über vergleichende satzmelodische Untersuchungen. Phonetica 5: 75–98 (1960).



Es sind drei Haupttypen melodischer Formen zu unterscheiden: ein abschließender, ein weiterweisender und ein interrogativer Typus. Das ist aber nicht die Terminologie des Verfassers. *Romportl* spricht vielmehr von den melodischen Formen der abgeschlossenen Nichtfragesätze, der abgeschlossenen Fragesätze (gemeint sind Entscheidungsfragen) und der Abschnitte vor einer nichtabschließenden Pause. Entsprechend ist auch das Kapitel «Melodie» gegliedert: A. «Der abgeschlossene Nichtfragesatz»; B. «Der abgeschlossene Fragesatz»; C. «Der Abschnitt vor einer nichtabschließenden Pause». (Es folgt dann noch: D. «Abriß des melodischen Systems der Teschener Dialekte».) Die Vor- und Nachteile der angeführten Gliederung zu behandeln und sich mit den verschiedenen Aspekten auseinanderzusetzen, die in ihr teilweise zur Geltung kommen, würden den Rahmen der Besprechung sprengen. Es darf nur gesagt werden, daß im Hintergrund der Gliederung (wenn auch unausgesprochen) die drei Haupttypen melodischer Formen stehen. Die drei «Satz»-Kategorien beruhen auf melodischen Kriterien; jede dieser Kategorien ist mit dem Anwendungsbereich eines der Haupttypen identisch.

Der abschließende Typus gelangt in Aussagen, Aufforderungen, Ergänzungsfragen und Ausrufen zur Anwendung. (Die Differenzierung dieser Subkategorien erfolgt größtenteils durch andere als melodische Mittel.) Der neutrale Satz wird durch eine konvex absinkende melodische Form gekennzeichnet, die mit der Akzentsilbe des sinnwichtigsten Wortes (des «psychologischen Prädikats») beginnt. Im emotionalen Satz wird dieser Verlauf abgeändert in eine tieftönig ebene oder steigend-fallende oder zirkumflekterte Melodieführung, die neben der Abgeschlossenheit gleichzeitig verschiedene emotionelle Gehalte signalisieren. Die weiterweisende Tonführung der nichtabschließenden Sätze zeichnet sich in einem neutralen Satz durch einen horizontalen Verlauf in der relativ mittleren Stimmlage aus, bei emotionaler Färbung dagegen durch einen steigenden Verlauf. Auch zum interrogativen Typus, der in Entscheidungsfragen angewendet wird, zählen mehrere Formen mit einem nichtsinkenden Verlauf, dessen steigende Tendenz besonders in der bedeutenden Erhöhung der Schlußsilbe zur Geltung kommt. Für die Anzeige emotionaler Gehalte ist die Lage der Akzentsilbe des «psychologischen Prädikats» bzw. auch die Lage der vorangehenden unbetonten Silbe als relevant anzusehen.

Man sieht, daß eine der Funktionen der 1. und 2. Ebene von der Melodie immer ausgeübt wird. Entweder Abgeschlossenheit oder Nichtabgeschlossenheit oder Fragebedeutung (Entscheidungsfrage) werden signalisiert. Die Funktion der 3. Ebene tritt dann nach Bedarf hinzu. Bedenkt man nun, daß zu einem der Haupttypen jeweils mehrere melodische Formen mit teilweise voneinander stark abweichenden Tonverläufen zählen, dann wäre doch eigentlich die Frage naheliegend gewesen, welche Merkmale den verschiedenen Formen des Haupttyps gemeinsam sind, durch welches Merkmal (oder welche Merkmale) sich die Formen, die neben der Grundfunktion des Haupttyps noch eine emotionelle Funktion ausüben, von jener unterscheiden, die in der 3. Ebene funktionell irrelevant ist, sowie durch welche Merkmale sich erstere noch voneinander differenzieren. Diese Fragen wurden vom Verf. bedauerlicherweise nicht gestellt. *Romportl* spricht zwar davon, daß eine melodische Form «durch ihre Gestaltung, genauer gesagt durch ihre relevanten Züge, die sich von den Zügen anderer melodischer Formen desselben mundartlichen melodischen Systems unterscheiden», die Bedeutung und Geltung (?) eines Satzes bestimmt (S. 76), doch deutet die Formulierung «von den Zügen anderer Formen desselben... Systems» darauf hin, daß er damit nicht die relevanten Merkmale im Sinne der obigen Fragen meinte. Das geht auch an mehreren Stellen aus seinen Ausführungen hervor. Nur im Bereich der interrogativen melodischen Formen nimmt Verf. einen kurzen Anlauf zu einer phonologischen Analyse. Der «Abriß des melodischen Systems der Teschener Dialekte» ist eigentlich eine kurze und übersichtliche Zusammenstellung des nach den dreierwähnten «Satz»-Kategorien angeordneten Inventars der in den untersuchten Mundarten automatisierten melodischen Formen.

In ihren allgemeinen Zügen stimmen die satzmelodischen Verhältnisse der Tschecher Mundarten im wesentlichen mit denen des Tschechischen überein. Abgesehen von Unterschieden im Inventar der melodischen Formen bzw. in der funktionellen Anwendung relativ gleicher Formen (diese Unterschiede wurden nicht systematisch behandelt) ist im Vergleich mit dem Tschechischen eine Besonderheit durch die VorletztSilbenbetonung gegeben. Im Normalfall, d. h. wenn das sinnwichtigste Wort am Satzende steht, bleibt der funktionstragende Teil der melodischen Linie des Satzes (die «melodische Form» oder «Kadenz» – der letztere Terminus wird vom Verf. nicht angewendet) auf die vorletzte und letzte Silbe eines Satzes beschränkt, während im Tschechischen die Länge infolge der Anfangsilbenbetonung variiert. Die Begrenzung kann nur infolge einer Akzentverschiebung im sinnwichtigsten Wort oder im Zusammenhang mit einer Veränderung der Struktur des Satzes (Umstellung des «psychologischen Prädikats») aufgehoben werden.

*Sprechtempo* und *Redepause* nehmen ebenfalls als «Sprechmodulationen» an der «Abgrenzung» bzw. an der «Bestimmung» (und wir fügen hinzu: auch an der Konstituierung) der Redeeinheit «Satz» teil. Während das Durchschnittstempo nur als charakteristischer Zug der Redeweise eines Sprechers zu betrachten ist (es gehört also in die sogenannte diagnostische Sphäre), können Tempoveränderungen innerhalb eines Satzes sprachlichen Wert erlangen. Die Pausen partizipieren an der Abgrenzung eines Satzes, doch kann diese delimitative Funktion nicht selbständig, sondern nur im Zusammenhang mit den anderen Faktoren ausgeübt werden.

In der *Zusammenfassung* stellt Verf. fest, daß die behandelten «Sprechmodulationen» in verschiedenem Maße und auf verschiedene Weise an der Abgrenzung und Bestimmung eines Satzes bzw. Satzabschnitts beteiligt sind. An der Abgrenzung partizipieren die nachstehenden Faktoren in folgender hierarchischer Abstufung: (1) Pause, (2) Dehnung der Endsilbe, (3) Vorhandensein einer melodischen Form. Diese Abstufung wird jedoch dadurch in Frage gestellt, daß *Romportl* beim Fehlen eines dieser Faktoren das Vorhandensein der übrigen als ausreichend erachtet, wobei als Beispiel eines fehlenden Faktors gerade die Pause (also das in der Abstufung des Verf. wichtigste Element) angeführt wird. Und im Kapitel über *Sprechtempo* und *Redepause* lesen wir, daß eine Pause ohne das Vorhandensein einer melodischen Form nicht als Mittel der Gliederung der zusammenhängenden Rede anzusehen sei (S. 85).

Die Bedeutung bzw. der «Charakter» eines Satzes oder Satzabschnitts wird vor allem durch die melodische Form bestimmt; hinzu treten Akzent sowie (hauptsächlich in der emotionellen Ebene) Quantitätsmodifizierungen.

Im ganzen glauben wir, daß mit *Romportls* Arbeit ein wertvoller Beitrag zur Erkenntnis der satzphonetischen Problematik im westslavischen Raum geleistet wurde. Kann man der Darstellung der Gegebenheiten ohne größere Einschränkungen voll zustimmen, so müssen doch gegen *Romportls* Ansichten über die Deutungsmöglichkeiten des Materials unter dem diachronischen Gesichtspunkt Bedenken angemeldet werden. Verf. möchte seine Arbeit nämlich nicht nur als phonetische Studie, sondern darüber hinaus als phonetischen Beitrag zur Lösung dialektologischer und linguistischer Fragen aufgefaßt wissen (S. 7 und im deutschen Résumé S. 110). Als eine dieser Fragen wird in der Zusammenfassung das Problem des Prozesses der Akzentstabilisierung im Westslavischen aufgeworfen. Angesichts der Gesamtsituation der Akzentuierung in den untersuchten Dialekten (stärkere oder geringere Neigung zu Akzentverschiebungen) und im Hinblick auf die Tatsache, daß Dreisilbler leichter als Vier- oder Fünfsilbler zu einer Akzentverschiebung auf die erste Silbe neigen, wird von *Romportl* die Vermutung ausgesprochen, daß die von ihm vorausgesetzte Entwicklung vom VorletztSilben- zum Anfangsilbenakzent (vom «polnischen» zum «tschechischen» Typus)<sup>18</sup> sich nicht auf

<sup>18</sup> Vgl. *M. Romportl*: K otázce ustálení západoslovanského přízvuku («Zur Frage der Stabilisierung des westslavischen Akzents»). Studie a práce lingvistické I (*Havránek-Festschrift*), S. 73–80 (Prag 1954).

dem Wege über einen Nebenakzent auf der Anfangssilbe viersilbiger Wörter (bzw. Akzenteinheiten) vollzogen habe, sondern eher von einer durch die Tendenz zur dynamischen Rhythmisierung des Satzes bedingten Akzentverschiebung in Dreisilblern ausgegangen sei. Beeinflußt von einer starken Frequenz dreisilbiger Wörter mit Anfangssilbenakzent könnte der durch die Tendenz zur dynamischen Rhythmisierung fast regelmäßig vorhandene Nebenakzent auf der ersten Silbe der Viersilbler sich sodann zum Hauptakzent entwickelt haben. Erscheint diese Hypothese mit Vorbehalten (besonders hinsichtlich des für derartig weittragende Schlußfolgerungen doch relativ beschränkten Materials) noch möglich<sup>19</sup>, so muß der Versuch abgelehnt werden, auf Grund einer Auslegung der gegenwärtigen Akzentverhältnisse hypothetische Erklärungen über den Ursprung der Bevölkerung abgeben zu wollen. Der bestimmte Grad der Labilität der Akzentuierung in der zusammenhängenden Rede wird vom Verf. dahingehend interpretiert, daß es in der Vergangenheit zu einer Mischung zweier Bevölkerungsgruppen kam, die zwei verschiedenen Akzentuierungstypen angehörten, und das höchstwahrscheinlich zu einer Zeit, als die Stabilisierung des Akzents auf der ersten bzw. vorletzten Silbe bereits durchgeführt war. Die Mundart mit der Pänultima-betonung habe zwar die Oberhand gewonnen, doch sei eine starke Neigung zur Akzentverschiebung geblieben. Damit dürften die Grenzen der Deutungsmöglichkeit des synchronischen Materials sichtlich überschritten worden sein.

Um abzuschließen: Mit der sehr sorgfältigen und präzisen Darstellung der Morphologie der satzphonetischen Elemente in den untersuchten Mundarten hat *Romportl* Vorzügliches geleistet. Die Arbeit enthält wertvolles Material und einige wichtige Schlußfolgerungen. Dadurch dürfte sie auch für zukünftige vergleichende satzphonetische Studien im Slavischen nicht ohne Bedeutung sein. Die Einwände, die im Verlauf der Besprechung erhoben wurden, stellen andere Ansichten, andere wissenschaftliche Meinungen dar und setzen den Wert der auch in den allgemein satzphonetischen Aspekten interessanten Arbeit nicht herab.

Hans-Walter Wodarz, Bonn

<sup>19</sup> Mit dem Für und Wider dieser Hypothese gedenken wir uns an anderer Stelle ausführlich zu befassen.

*Kita Tschenkeli: Einführung in die georgische Sprache.* Band I, Theoretischer Teil, LXIV + 628 Seiten. Band II, Praktischer Teil, X + 614 Seiten. Amirani Verlag, Zürich 1958.

La parution d'une grammaire de la langue géorgienne dans une langue européenne est un événement qui mérite qu'on s'y arrête. Les moyens dont on a disposé jusqu'à ces temps derniers pour l'étude de cette langue sont, en effet, très insuffisants. Les grammaires de *A. Šanize*, écrites en géorgien, ne peuvent, pour répandre la connaissance du géorgien en Europe, rendre que des services très limités, comme elles supposent connue la langue qu'elles décrivent. La petite grammaire d'*Adolf Dirr* qui, pendant des années était la seule à notre disposition, était très mauvaise à tous les égards. Celle de *Rudenko*, publiée en russe en 1940, avait des qualités, mais dans son classement des formes verbales, la seule partie vraiment compliquée de la langue, était trop marquée de l'influence de *Marr* pour être tout à fait satisfaisante\*, et mon Esquisse d'une grammaire du géorgien moderne, Oslo 1936, dont je me garderai de dire du bien ou du mal, est actuellement épuisée. Le présent ouvrage comble ainsi une lacune, et il le comble d'une façon remarquable.

Deux magnifiques volumes de près de 1100 pages, très soigneusement imprimés

\* Voir mon compte-rendu dans *Norsk Tidsskrift for Sprogvidenskap*, t. XII, p. 272-276, Oslo 1942.



et solidement reliés, l'un contenant une description très complète de la morphologie, l'autre des exercices, une chrestomathie et des notes très copieuses avec vocabulaire – le tout pour 68 francs suisses – c'est là un beau monument que l'auteur a dressé en l'honneur de sa langue maternelle.

Dans l'introduction l'auteur donne un bref aperçu du peuple géorgien et de son histoire et cite les vues d'un grand nombre de savants sur l'importance des études géorgiennes. L'auteur nous informe en grand détail de toutes les difficultés qu'il a eu à vaincre et finit par nous expliquer le choix de sa méthode de présentation. Sur ce dernier point j'aurai à revenir. Pour ma part j'aurais souhaité une attitude plus nette de la part de l'auteur vis-à-vis des théories de *Marr*. Plutôt que de signaler à l'attention du lecteur la traduction allemande d'un de ses travaux, il aurait mieux valu les passer sous silence. Elles ont à un certain moment menacé de discréditer les études caucasiennes\*. Dans la partie phonétique, incorporée d'une façon assez surprenante dans l'introduction, l'auteur signale l'importance des groupes consonantiques si caractéristiques du géorgien et consacre quelques pages à l'exposé des groupes qu'il appelle «harmoniques», c.-à-d. composés de termes ou sonores ou sourds aspirés ou glottalisés. Si ces groupes sont mentionnés, on s'attend à quelques mots sur les groupes assez nombreux qui n'entrent point dans cette catégorie. Ils semblent avoir échappé à l'attention de l'auteur, bien que leur existence et les problèmes de classement qu'ils soulèvent soient mentionnés dans mon Esquisse. On regrette que l'auteur n'ait pas pu utiliser les importants travaux de *Sergi Žyent'i* sur la phonétique géorgienne, en particulier sur l'accentuation.

Dans la grammaire proprement dite l'auteur s'est laissé guider par des considérations d'ordre pédagogique. Il a renoncé de parti pris à la description systématique des différentes classes de formes et de mots, pour répartir toute la matière sur 48 leçons qui se succèdent dans l'ordre de complexité croissante. Dans la description du système nominal l'auteur nous donne quelques formes verbales très fréquentes, ce qui lui permet dès les premiers exercices du 2<sup>e</sup> volume de présenter un certain nombre de petites phrases et d'enrichir le vocabulaire du lecteur de leçon à leçon. Cette méthode à laquelle l'auteur semble beaucoup tenir, peut avoir ses avantages, mais elle a aussi ses inconvénients. L'unité de chaque partie de la morphologie risque d'être perdue de vue dans ce morcellement de la description. Le système des cas est ainsi exposé dans la première leçon, le cas ergatif n'est expliqué que dans la 18<sup>e</sup>. Les préfixes pronominaux subjectifs sont donnés dans la 8<sup>e</sup> leçon, les préfixes objectifs dans la 31<sup>e</sup>, et ainsi de suite. Cette méthode par leçons savamment dosées est bien connue des manuels dits pratiques. Les linguistes qui s'intéressent surtout aux structures l'aimeront moins bien. Heureusement la table des matières est assez détaillée, de sorte que le linguiste pourra assez facilement, en s'en servant, rassembler les membres disloqués pour en reconstituer l'unité.

La grammaire est extrêmement bien documentée, l'exposé est clair avec une profusion de paradigmes et des envois très nombreux d'une leçon aux autres. Il n'appartient évidemment pas à un étranger de porter un jugement sur la correction idiomatique des exemples donnés, les Géorgiens auront à se prononcer là-dessus. Mais j'ai l'impression que l'auteur est partout un guide sûr, qui a beaucoup réfléchi sur sa langue et qui a su mettre à profit la plus grande partie de la littérature déjà existante sur le géorgien. Il y a peu de problèmes qui n'aient pas reçu un traitement dans une partie ou une autre de l'ouvrage. Il est vrai que la syntaxe n'a pas été traitée à part, mais la plupart des faits syntaxiques sont étudiés avec la morphologie. On aurait souhaité quelques pages au moins sur l'ordre des mots dans la phrase.

On doit louer l'auteur de se baser partout sur des faits purement formels dans sa description des formes, fait assez rare dans les grammaires qui visent un public plus

---

\* On peut maintenant renvoyer à l'ouvrage très consciencieux de *Lawrence L. Thomas*, *The Linguistic Theories of N. Ja. Marr*, Berkeley/Los Angeles 1957, dont j'ai donné un compte-rendu dans *Romance Philology* vol. XII, 1, Berkeley 1958.



large que les spécialistes. Il y a pourtant une exception : l'établissement de la classe des verbes indirects a été fait sur des critères sémantiques. C'est la classe des verbes tels que ,j'ai sommeil, je crains, j'aime, j'oublie', etc. où la personne qui est le siège de ces sentiments apparaît en géorgien comme régime indirect et où ce qui, en français ou en allemand, est régime direct, apparaît en géorgien comme sujet grammatical. Cette procédure est peu recommandable, car elle signifie qu'on classe des formes géorgiennes d'après leur traduction dans une autre langue. Or, il est parfaitement possible de procéder ici aussi avec des critères formels, si l'on part de l'expression du pluriel. La classe des verbes moyens est par contre très bien définie, bien que, à mon avis, elle ne puisse pas être mise sur le même plan que les classes des verbes actifs et passifs. C'est une classe qui recoupe les deux classes mentionnées, étant composée d'une partie des verbes actifs et d'une partie des verbes passifs. Mais tout cela n'a au fond qu'un intérêt théorique et on comprend que l'auteur n'y ait pas trop insisté.

Le classement des formes verbales, qui est la partie la plus difficile de toute description du géorgien, est en général très bien fait. Il ne semble pas se distinguer de celle de *Sanize* ni de celle que j'ai appliquée dans mon *Esquisse*. Mais sur beaucoup de points l'auteur apporte des précisions excellentes et parfois il ajoute à notre connaissance des faits. Je voudrais surtout attirer l'attention sur l'exposé des aspects et du rôle des préverbes dans les différentes classes des verbes. La «polyvalence» de beaucoup de formes verbales est très bien mise en relief, comprenant par ce terme le fait que la valeur d'une forme verbale est ambiguë tant que l'on ne sait pas dans quelle série de formes elle s'insère. Ainsi la forme *ec'ereba* est polyvalente : dans la mesure où elle s'oppose à *ic'ereba* c'est le présent du passif relatif de procès du verbe *c'era* écrire. Dans la mesure où elle s'oppose à *sc'eria*, *ac'eria* ou *uc'eria* c'est le futur du passif d'état du même verbe, et ainsi dans beaucoup de cas.

L'ouvrage a deux défauts. Dans son désir très louable de tout dire, l'auteur est tombé dans une prolixité qui gêne depuis l'introduction jusqu'aux dernières pages. La présentation d'un ensemble de formes dans des paradigmes, c'est excellent, mais la multiplication de paradigmes parfaitement réguliers, absolument identiques au point de vue formel, c'est pour le lecteur une source d'obscurité plutôt qu'un avantage. Les combinaisons des préfixes subjectifs et objectifs dans une seule et même forme, c'est évidemment une difficulté pour beaucoup de lecteurs, mais dans le présent ouvrage l'exposé de ces combinaisons remplit des pages et des pages, répété pour chaque classe verbale. Un effort de concision et un emploi modéré de quelques tableaux d'ensemble auraient permis de réduire le nombre des pages considérablement, d'un tiers peut-être. C'est là une économie dont le livre aurait profité à tous égards et elle aurait permis de remédier au deuxième défaut de l'ouvrage, l'absence d'un index. Le lecteur qui cherche les formes de futur des présents tels que *všvreb* je fais, *vambob* je dis, *vgržnob* je sens, *vnneb* je nuis, se trouvera bien embarrassé. Je suis sûr que l'auteur saurait nous dire où l'on trouverait ces renseignements, soit dans la grammaire soit dans le deuxième tome. Avec un index d'une vingtaine ou d'une trentaine de pages le lecteur les aurait trouvés sans difficultés.

J'ai peu de remarques à faire à propos du deuxième volume que je n'ai pas pu examiner en détail. Il contient 394 pages d'exercices de 48 leçons de grammaire et 217 pages de textes littéraires, tous pourvus d'une traduction en allemand et de notes explicatives très détaillées. Tous les matériaux lexicaux de ce volume constitueraient à eux seuls un petit dictionnaire, et on finit par se demander s'il n'aurait pas mieux valu les présenter en appendice, dans l'ordre alphabétique.

Les critiques qu'on peut adresser à l'auteur sont toutes d'importance secondaire, elles ne touchent pas au fond. Nous avons maintenant en main un outil perfectionné, bien que de maniement un peu lourd, pour l'apprentissage et l'enseignement du géorgien. Nous n'avons rien de semblable pour aucune autre langue du Caucase.

Hans Vogt, Oslo

**Felix Kahn: Phonétique et grammaire comparatives pour l'enseignement et l'allemand dans les écoles primaires et secondaires de langue française.** Cahiers Ferdinand de Saussure 16, Librairie E. Droz, Genf 1958. 90 S.

Das Interesse, das *Felix Kahn* Strukturproblemen der modernen Sprachen entgegenbringt, hat er bereits durch seine Arbeit «Le système des temps de l'indicatif chez un Parisien et chez une Bâloise» (Genf 1944) dokumentiert. Freilich, wie sich noch zeigen wird, ist es etwas anderes, ob sich eine derartige Untersuchung auf die Muttersprache oder eine Fremdsprache bezieht.

Behandelt werden: I. Faits phoniques allemands sans équivalent exact en français. II. Rapports entre la prononciation et l'orthographe. III. Fait grammaticaux allemands sans équivalent exact en français (A. Genres, B. Cas et prépositions, «zu» devant l'infinif, conjunctions, adverbes, C. La place des termes).

*Kahn* setzte sich diesmal ein vorwiegend didaktisches Ziel: eine vergleichende Lautlehre und Grammatik des Deutschen und Französischen zu Unterrichtszwecken. Allerdings ist der Grundsatz «le maître ne s'attardera pas longtemps aux faits de la langue étrangère qui sont semblables à ceux de la langue maternelle» diskutabel. Gerade Ähnlichkeiten täuschen oft über sehr störende Verschiedenheiten hinweg; so ist es, um nur ein Beispiel zu nennen, völlig abwegig, das «e» des Wortes «Küste» mit dem sogenannten «e muet» artikulatorisch gleichzusetzen. Der deutsche Laut ist ungerundet, der französische gerundet! Zweifellos eine häufige Fehlerquelle des deutsch sprechenden Franzosen.

Zu S. 37: Da die Bühnensprache nur ein mittleres «a» kennt, so unterliegt dieser Vokal nicht in derselben Weise dem Gegensatz: langer geschlossener – kurzer offener Vokal.

Sehr zu begrüßen ist der Hinweis auf das noch nicht genügend bekannte Unterbleiben der regressiven Assimilation (z. B. Innsbruck) entsprechend französischen Fällen wie: une patte d'oie, wo Sonorisierung der Tenuis eintritt (41).

Zur Grammatik: Die didaktische Bestimmung des Buches verbietet es grundsätzlich, entwicklungsgeschichtlich bedingte Schwankungen in besonderem Maße zu berücksichtigen. In allen Fällen über sie hinwegzugehen, halte ich jedoch nicht für ratsam. Beim Geschlecht der französischen Städtenamen etwa: mit Recht bemerkt *Dauzat* (dessen ausgezeichnete «Grammaire raisonnée de la langue française» leider nicht benutzt wurde) mit Bezug auf die Städtenamen, die auf ein «e instable» ausgehen: «Tout au moins pour les villes françaises, le masculin l'emporte» (p. 79).

Es wäre von Vorteil gewesen, darauf hinzuweisen, daß deutsche Länder- und Gebietsnamen auf -ci, -ie oder -e weiblich sind (Duden 143). «Außer» ist im Gegensatz zur Auffassung *Kahns* eine durchaus gebräuchliche Präposition (56). Bei «trotz» steht neben dem Genitiv auch der Dativ (57), ebenso bei «längs». Bei der Frage, wann «de» oder «par» beim Urheber der passiven Handlung steht, genügt es nicht, nur auf *Grevisse* hinzuweisen; *Dauzat*, Gr.r.355 gibt diesbezüglich wertvolle zusätzliche Bemerkungen.

Beim Genus verbi (58 ff.) wäre auch der Typus «Défense de parler ici» zu erwähnen gewesen. Ebenso: «Il s'est vu refuser le visa.»

Es ist irreführend, zu behaupten (64), «l'allemand emploie l'imparfait plus souvent que le français». So zu argumentieren, wird erst möglich, wenn man willkürlich die «savais»-Form mit der «wußte»-Form unter einer Etikette zusammenfaßt. Außerdem könnte man dem Beispiel: «Ich kam, sah und siegte» (je suis venu usw.) andere entgegenstellen, z. B. «hast du das gewußt?» – «tu le savais?».

Bei einer vergleichenden Strukturlehre des Deutschen und Französischen müßte auch auf die Wiedergabe deutscher Participia praesentia hingewiesen werden. Sehr häufig entspricht im Französischen eine andere Ausdrucksform: Relativsatz (Das Bild stellt ein schlafendes Kind dar: ... un enfant qui dort), Part. perf. (un enfant endormi),

Präposition + Subst. (Fliehende Truppen: des troupes en fuite). Schon *Dubray* wies darauf hin, daß demgegenüber das französische Particip praes. häufig adjektivischen Charakter hat (un visage riant = ein heiteres Gesicht; brûlant = brennend heiß usw.).

Die Arbeit *Kahns* ist ein sehr verdienstvolles Bemühen, Strukturunterschiede des Deutschen und Französischen in einer ansprechenden lebendigen Art aufzuzeigen. Leider verliert das Buch dadurch ungemein an Wert, daß es nicht in Zusammenarbeit mit einem Deutschen verfaßt beziehungsweise von einem solchen kontrolliert worden ist\*. Wenn *Kahn* schreibt: «La plupart des mots sont forgés pour les besoins de la cause, mais d'après des modèles qui me sont familiers depuis des années» (36), so ist diese «familiarité» manchmal zu bedauern. Eine Kostprobe: «Gestern bin ich weder zu Hause geblieben, noch zu meiner Großmutter gegangen. Zum Geburtstag und zum Ausruhen (!) durfte ich um den See fahren. Ohne Angst (!) nahmen wir statt eines Dampfers ein kleines Motorboot. Gestern hatte ich die Erlaubnis, auf den See zu gehen (!), um einen schönen Geburtstag zu haben und mich richtig auszuruhen. Ohne uns vor Schwierigkeiten zu fürchten (!), nahmen wir ein kleines Motorboot, statt auf einem Dampfer zu fahren. Wir fuhren nicht zu schnell, aber unsere Reise wurde zu lang (!). Als wir heimkamen, waren sogar die Bäckereien zu, und manche Leute machten schon die Fensterläden zu.»

Oder (S. 76): «Wovon spricht Hans, vom Wetter? Ja, das ist es, wovon er immer spricht (!). Woran denkst du, an deinen Aufsatz? Nein, es ist nicht das, woran ich denke (!).» Hier wird es besonders deutlich, wie *Kahn* mit deutschen Wörtern französisch zu sprechen versucht. Ebenso, wenn es heißt: «Die Kranke wird vom Arzt operiert. Sie ist von einem Sturm erschüttert (!! ) und durch einen Stein verletzt worden.» Es ließen sich noch zahlreiche Beispiele dieser Art anführen, die das Buch, solange es nicht von einem Deutschen korrigiert wird, für den Unterrichtsgebrauch untauglich machen.

Wolfgang Pollak, Wien

\* Etwa nach dem Vorbild der ausgezeichneten «Grammaire de l'allemand» von J. Fourquet (Paris 1952).

*André Martinet: Eléments de linguistique générale.* Librairie Armand Colin, Paris 1960 (Collection Armand Colin, Section de littérature. 349.) 224 p.

*André Martinet*, «maestro y guía de una nueva generación de linguistas»<sup>1</sup>, legt die Grundzüge der allgemeinen Sprachwissenschaft dar, wie er sie in ähnlicher Form seinen Hörern an der Sorbonne vorträgt. Dies ist allerdings ein Buch «für Leser aller Fakultäten»: Auseinandersetzungen sind selten, Belege und Zitate fehlen ganz, und nur ein kurzes, geschickt ausgewähltes Literaturverzeichnis (S. 218–220) führt in das Schrifttum ein. Der fachfremde Leser mag daher gelegentlich eine plausible Hypothese für ein gesichertes Ergebnis halten; er wird die These kennenlernen, doch nicht die Antithese, und die Frage nach der Möglichkeit einer Synthese wird ihm nicht in den Sinn kommen.

Solche Nachteile sind bewußt in Kauf genommen (S. 6) und mindern das Verdienst des Buches nicht. Wir sehen mindestens zwei Gründe, es willkommen zu heißen: 1. Jeder Interessierte erhält die Chance, sich über einen Teil der Methoden zu unterrichten, die die moderne Sprachwissenschaft charakterisieren; 2. Linguisten, Philologen, Phonetiker und nicht zuletzt die Studenten der entsprechenden Fächer haben die Auffassungen eines der einflußreichsten Linguisten unserer Tage in zusammenhängender Form bequem zur Hand (obgleich man in Einzelfragen auf die zahlreichen Spezialarbeiten des Verfassers zurückgreifen muß)<sup>2</sup>. Dies aber scheint einem dringenden Bedürfnis entgegenzukommen. Kaum ein Linguist beherrscht heute das Gesamtgebiet der

<sup>1</sup> Estructuralismo e historia. Miscelánea homenaje a André Martinet (Bd. 1–2. La Laguna 1957. Mit Bibliographie bis z. J. 1956), Bd. 1, S. 5.

<sup>2</sup> Zur Bibliographie vgl. Anm. 1.



Linguistik; es ist fast unmöglich geworden, alle Probleme und ihre Lösungsversuche im Auge zu behalten. Überdies unterscheiden sich verschiedene Schulen in ihren methodologischen Annahmen, und die Entwicklung der Terminologie scheint zentrifugalen Kräften zu gehorchen – konstant aber bleibt die menschliche Arbeitskraft. Solange zu ihrer Entlastung nicht auch auf unserem Gebiet elektronische Hilfsmittel eingesetzt werden können, gehören methodologisch orientierende Zusammenfassungen ebenso zu unserem wissenschaftlichen Existenzminimum wie Bibliographien, terminologische Wörterbücher und die private Spezialkartei.

In Deutschland werden Sprachwissenschaft und Sprachphilosophie oft verwechselt; Betrachtungen über das «Wesen» der Sprache können als Einführung in die allgemeine Sprachwissenschaft ausgegeben werden. Deshalb sei *Martinets* Auffassung von Sprachwissenschaft und Sprache kurz wiedergegeben.

Für *Martinet* ist die Sprachwissenschaft eine *beschreibende* Wissenschaft, ihr Gegenstand primär die *gesprochene* Sprache als eine *Institution*, die eine Verständigungsgemeinschaft konstituiert. Sie erfüllt diese ihre Hauptfunktion durch eine zweifache Gliederung (*double articulation*)<sup>3</sup>:

1. Die Erfahrungsdaten der Sprachgemeinschaft werden so ausgewählt und geordnet, daß sich jede Situation durch eine bestimmte Anordnung einer endlichen Anzahl isolierbarer Einheiten darstellen läßt, denen je ein Lautkomplex zugeordnet ist (*première articulation*);

2. die Lautkomplexe ihrerseits sind aus Lauteinheiten aufgebaut, deren Anzahl begrenzt, aber groß genug ist, um in bestimmten Kombinationen jeden Lautkomplex im Prinzip unverwechselbar zu machen (*deuxième articulation*).

Beide Gliederungen schließen Abstraktionsvorgänge ein. Im Prinzip wird jede Sprache auf ihre eigene Weise abstrahieren. Eine Sprache ist daher keine Nomenklatur für vorgegebene Gegenstände (S. 14). Angesichts der in Theorie und Praxis immer wiederkehrenden Identifizierung von «Wort» und «Sache» kann dieses Prinzip nicht oft genug ausgesprochen werden. Nicht zuletzt liegt es auch der historisch-vergleichenden Methode zugrunde<sup>4</sup>.

Ein beliebiger Lautkomplex, der eine «Bedeutung» hat, wird ein Zeichen (*signe*, bestehend aus *signifiant* und *signifié*) genannt; die kleinsten Zeichen heißen Moneme, die kleinsten Einheiten der zweiten Artikulation sind die Phoneme.

Gegliedertheit der Sprache und Linearität des Sprechvorgangs zwingen den Sprecher an jedem Punkt des Sprechablaufs zu einer Wahl unter den an diesem Punkt zugelassenen Einheiten des Systems. Jede Einheit steht daher in einem doppelten Bezugsverhältnis: Sie steht in einer syntagmatischen Beziehung, d. h. in *Kontrast* zu ihren Nachbareinheiten; und in einer paradigmatischen, d. h. in *Opposition* zu den Einheiten, die in einer bestimmten Äußerung an ihrer Stelle stehen könnten<sup>5</sup>.

Dies sind die allgemeinen Merkmale, die ein Verständigungssystem nach *Martinet* aufweisen muß, wenn es zum Gegenstand der Sprachwissenschaft werden soll. Diese Merkmale sind per definitionem universell; was darüber hinaus universell sein mag, muß die Erforschung des Spezifischen in allen Einzelsprachen ergeben.

Wir haben die Definition deshalb so ausführlich wiedergegeben, weil ihre Bestandteile den Wert von Axiomen in *Martinets* Auffassung von Sprachwissenschaft haben,

<sup>3</sup> Vgl. *A. Martinet*: La double articulation linguistique. TCLC 5: 30–37 (1949) (Recherches structurales 1949); ders.: Arbitraire linguistique et double articulation. CFS 15: 105–116 (1957).

<sup>4</sup> *A. Meillet*: Introduction à l'étude comparative des langues indo-européennes. 8. Aufl. (hg. v. E. Benveniste, Paris 1937), S. 14–15.

<sup>5</sup> Vgl. auch *A. Martinet* in Word 9: 9, Anm. 21 (1953); ders. in Actes du VIII<sup>e</sup> Congrès International des Linguistes, Oslo 1958, S. 213; und ausführlicher *L. J. Prieto*: Traits oppositionnels et traits contrastifs. Word 10: 43–59 (1954).



unabhängig davon, ob sie Hypothesen oder Fakten sein mögen. Sie bilden ein Bezugssystem für die nachfolgenden Einzelheiten, und diese können nur aus jenen verstanden werden. So werden zunächst die Prinzipien der Sprachbeschreibung abgeleitet:

Eine Sprache kann als Verständigungssystem sachgemäß nur synchronisch beschrieben werden, obgleich sich jede Sprache verändert. Für die Beschreibung einer Einzelsprache ist nur das relevant, was der *Wahl* des Sprechers unterliegt, d. h. die Einheiten des Sprachsystems mit ihren Funktionen. Insbesondere kann nur *die* Information als übermittelt angesehen werden, die in irgendeiner Weise als formal manifestiert nachgewiesen werden kann (S. 42–43). Auf die Gefahren einer Analyse auf Grund von Übersetzungen wird ausdrücklich hingewiesen. Nur aus praktischen Gründen (S. 41 oben) schließt *Martinet* das Verfahren, durch eine Untersuchung der Distribution der *signifiants* allein zu einer Analyse des Zeichensystems zu gelangen, als unzumutbar aus. Es muß dazu angemerkt werden, daß gegen dieses Verfahren aber auch schwer zu widerlegende theoretische Argumente vorgebracht worden sind, vor allem der Vorwurf der Zirkularität<sup>6</sup>.

Es folgt S. 45–51 eine Aufführung der Grundbegriffe der artikulatorischen Phonetik mit der Begründung, sie biete nicht nur ein bequemes Verständigungsmittel, sondern auch «*permet de mieux percevoir la causalité des changements phonétiques*» (S. 45); wir müssen den zweiten Teil dieser Begründung den Phonetikern und Phonologen zur kritischen Prüfung überlassen und begnügen uns damit, vor einer Oberflächenbehandlung der Kausalität zu warnen – sicherlich im Sinne des Verfassers der «*Economie des changements phonétiques*»<sup>7</sup>.

Die Funktionen der Phoneme sind nach *Martinet* distinktiv = oppositiv, kontrastiv und expressiv (S. 52–53). Wir möchten vorschlagen, den Gebrauch der Ausdrücke «*oppositiv*» und «*kontrastiv*» auf die Beziehungen der Phoneme (und der Moneme) untereinander zu beschränken (so S. 33!, vgl. hier oben S. 185), dagegen die Funktionen der Phoneme in bezug auf die Identifikation bzw. die Abgrenzung von Monemen als distinktiv bzw. demarkativ (so auch *Martinet* S. 91–93!) zu bezeichnen und eine Vermischung dieser Gebrauchsweisen zu vermeiden.

Die Aufstellung eines Phonemsystems wird eingehend dargestellt: Zunächst werden Segmente in ihre kleinsten Bestandteile, die Einzellaute, zerlegt; die Technik kann als bekannt vorausgesetzt werden. Als zweiter Schritt werden die Kontextbedingungen eliminiert. Während es nun gängige Praxis ist, davon auszugehen, daß «*phonetisch ähnliche*» Laute Allophone eines Phonems darstellen, schlägt *Martinet* eine differenziertere Methode vor: Für jeden Laut in bestimmter Umgebung wird festgestellt, in welchen Beziehungen er zu den übrigen Lauten steht, die in dieser Umgebung an seiner Stelle vorkommen könnten. Zur Definition der «*Beziehungen*» dienen die *traits pertinents*, die kleinsten lautlichen Merkmale, die hinreichen, um zwei Laute (einer gegebenen Sprache) voneinander zu unterscheiden. Die Beziehungen der Laute zueinander werden für jede lautliche Umgebung gesondert bestimmt. Ein Phonem stellen schließlich diejenigen Laute dar, die in jeder lautlichen Umgebung zu den übrigen Lauten, die an ihre Stelle treten könnten, in gleichartigen Beziehungen stehen<sup>8</sup>. Die Definition ist also paradigmatisch, nicht syntagmatisch (distributionell). Ergänzend

<sup>6</sup> H. Frei: *Critères de délimitation*. Word 10: 136–145 (1954); E. Haugen: *Directions in Modern Linguistics*. Lg. 27: 211–222 (1951); (zugl.): *Readings in Linguistics*. The development of descriptive linguistics in America since 1925, hg. v. M. Joos, 2. Aufl. New York 1958, S. 357–363.

<sup>7</sup> A. Martinet: *Economie des changements phonétiques*. Traité de phonologie diachronique. Bern (1955). (Bibliotheca Romanica. Series prima. 10.)

<sup>8</sup> Vgl. hierzu auch A. Martinet: *Substance phonique et traits distinctifs*. BSL 53, Fasc. 1: 72–85 (1957–1958).

werden die Begriffe Variante, Neutralisation und Archiphonem besprochen (S. 67–73)<sup>9</sup>.

Bei der Erörterung des Unterschiedes von Vokalen und Konsonanten (S. 74–75) wird als Spezialfall erwähnt, daß im Französischen *i* und *j* nicht ein Phonem (mit zwei Varianten) darstellten, sondern zwei Phoneme. Zur Begründung werden die Gegensatzpaare *pays:paye*, *abbaye:abeille* angeführt. Dazu heißt es ferner, daß die Opposition von *i* und *j* überall neutralisiert sei, außer im Silbenauslaut. Nun ist aber kurz vorher (S. 74) implizit zugestanden worden, daß *pays* zweisilbig ist. Wir sehen nicht, wie sich die beiden Aussagen miteinander vereinbaren lassen. Es scheint uns auch nicht glücklich, die Fälle vom Typ fr. /kao/:/kap/<sup>10</sup> als Beispiel dafür heranzuziehen, daß Vokale und Konsonanten in den gleichen Umgebungen vorkommen können.

Die Schwierigkeiten, die die prosodischen Erscheinungen bieten, sind bekannt. Im Rahmen der *Martinet*schen Grundannahmen stellen sie sich u. a. so dar: Die prosodischen Erscheinungen sind nicht der zweiten Artikulation unterworfen; in irgendeiner Weise sind Intonation, Akzent, Tonhöhe, Energie, Lautdauer stets beim Sprechvorgang anzutreffen, der Sprecher hat daher nicht die Wahl zwischen ihrer An- und Abwesenheit; «en conséquence, ils se prétent moins bien à caractériser des unités discrètes» (S. 77). Intonationskurven sind darüber hinaus prinzipiell nicht in diskrete Einheiten auflösbar, also nicht der doppelten Artikulation unterworfen, und haben daher nicht einen «caractère proprement linguistique» im Sinne von *Martinet*s Definition. Dem entspricht es, daß sie «des fonctions mal différenciées» haben (S. 79). Es gebe z. B. für das Französische keine deutliche Grenze zwischen Aussage- und Frageintonation. Uns scheint allerdings, daß der Hörer dennoch nur drei Möglichkeiten des Verstehens behält: (1) Aussage, (2) Frage, (3) Zweifel an der Wahl zwischen (1) und (2); dagegen gibt es kein drittes Signal, das er auffassen könnte. Expressive Nuancen können in (1) wie auch in (2) übermittelt werden. Sind also beim Sender keine Stufen aufweisbar, so doch beim Empfänger, und der Sender *kann* in jedem Falle deutliche Signale geben. Es mag somit auf die Methode ankommen, ob man diskrete Einheiten ermittelt oder nicht.

Das Kapitel «Les unités significatives» (S. 97–145) enthält die Ergebnisse eines Versuchs (vgl. S. 7), eine umfassende Theorie der Moneme aufzustellen<sup>11</sup>. Zunächst wird die Analyse von Äußerungen in kleinste bedeutungstragende Einheiten besprochen, nämlich die Moneme, ihre Varianten (/žvè/ «je vais»: /tūva/ «tu vas»; /žpō/ «je peux»: /žpūi/ «je puis») und ihre Verschmelzungen (*amalgames*): /o/ + C z. B. «au professeur»: /a l/ + V z. B. «à l'étudiant».

(Zu den Varianten des engl. nominalen Pluralmonems heißt es [S. 103]: «...l'alternance de /s/ et de /z/ n'est toutefois pas déterminée *phonologiquement* par le contexte phonique, puisque *since* /sɪnz/ existe à côté de /sɪnz/; son conditionnement peut être formulé en termes phoniques, mais il ne vaut que pour un emploi morphologique déterminé.» Eine solche morphologisch bedingte Sonderbehandlung von Phonemen mag aus verschiedenen Sprachen zu belegen sein; für das zitierte Beispiel gilt sie jedoch nur dann, wenn man /s/ als die Grundform des Monems ansetzen will; doch nichts zwingt uns dazu, und die Verteilung der Varianten /s/ ∞ z ∞ ɪz/ läßt sich zwanglos plausibel machen, wenn man /ɪz/ oder /z/ [mit unterschiedlichen Konsequenzen] als Grundform annimmt.)

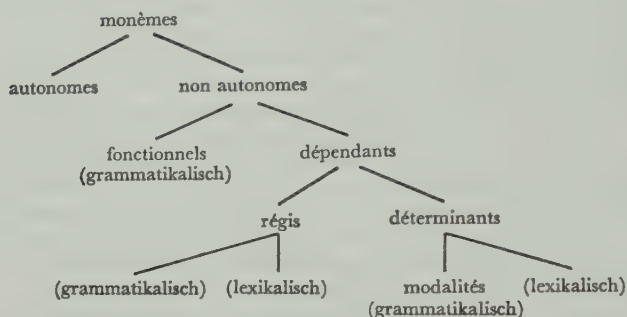
Anschließend (S. 104–127) werden verschiedene Unterscheidungen getroffen, die zu einer Hierarchie der Moneme führen sollen. Unterschieden werden drei Arten, eine

<sup>9</sup> Anwendungsbeispiele finden sich bei *A. Martinet*: La description phonologique, avec application au parler franco-provençal d'Hauteville (Savoie), Genf 1956, sowie bei *Don Graham Stuart* und *Matthew M. Haldol*: The Phonology of the Word in Modern Standard Mongolian. Word 13: 65–98 (1957).

<sup>10</sup> Ebenso z. B. in: Substance phonique... (vgl. Anm. 8), S. 72–73.

<sup>11</sup> Vgl. *A. Martinet*: Elements of a Functional Syntax. Word 16: 1–10 (1960).

Funktion zu kennzeichnen, zwei Typen von Funktionen (primär und sekundär) und danach verschiedene Arten von Monemen: autonome, funktionelle und «abhängige», die wieder in «régis» und «déterminants» zerfallen; ein anderes Einteilungsprinzip ergibt grammatische Moneme (= Morpheme) und lexikalische Moneme (= Lexeme, früher Semanteme genannt); die grammatischen *monèmes déterminants* werden *modalités* (engl. modifiers) genannt (Artikel, Genus-, Numerusformantien u. dgl.). Es ergibt sich somit, wenn wir recht verstanden haben, folgendes Bild:



Man wird finden, daß diese Einteilung und der zugrunde liegende Text mehr als ein Problem enthalten, das «neuf, trop neuf pour un manuel comme celui-ci» (vgl. S. 7) ist. Um einige davon zu nennen: In welchem Sinne kann man sagen, daß «il y a des énoncés tels que *je partirai demain* dont le sens ne varie jamais si je change l'ordre de certains monèmes et dis *demain, je partirai*» (S. 105)? – Ist die angegebene Definition von «Funktion» methodisch fruchtbar? Die Formulierung «*Fonction désigne ici le fait linguistique qui correspond au rapport entre un élément d'expérience et l'expérience globale*» (S. 109) scheint uns nicht ohne Zirkularität anwendbar. – Können alle Abhängigkeitsverhältnisse unter dem Stichwort «Subordination» zusammengefaßt werden? – Ist es hinreichend gesichert, daß die *monèmes autonomes* wie franz. *hier, demain, aujourd'hui* «impliquent leur rapport avec le reste de l'énoncé» (S. 107) und daß in einem «autonomen Syntagma» wie *l'an dernier* «c'est l'ensemble des monèmes en cause qui indique son rapport avec le contexte» (S. 109)? – Verbirgt sich hinter der Feststellung «le déterminant grammatical *le* est un élément centripète, le monème fonctionnel avec un élément centrifuge, selon le schéma  $\leftarrow$  avec le  $\rightarrow$  *sourire*» (S. 118) der Ansatz für ein neues Modell zur Beschreibung syntaktischer Konstruktionen?

Die Unterscheidung von *monèmes fonctionnels* und *modalités* auch und gerade dort, wo sie miteinander verschmelzen oder sich morphologisch ähnlich verhalten, scheint und besonders bemerkenswert; für viele hat sie ungewohnte Konsequenzen für die Beschreibung von Genus, Numerus, Kongruenz u. a. m. (sie ist aber implizit z. B. bereits von Harris unter dem Stichwort «morphemic long components» vorgenommen worden)<sup>12</sup>. Im übrigen wird man abwarten müssen, wie weit dieser Entwurf zu einer funktionellen Syntax ausgebaut werden muß, wenn er auf die Beschreibung einer konkreten Sprache angewendet wird; möglicherweise muß er durch eine andere Konzeption ergänzt werden.

Das 5. Kapitel ist der Vielfalt der sprachlichen Grenzübergänge gewidmet; der Begriff der Einzelsprache wird fließend; altersmäßige, soziale, geographische Unterschiede erhöhen die Schwierigkeit, zwischen Dialekt und Sprache scharfe Grenzen zu

<sup>12</sup> Zellig S. Harris: *Structural Linguistics*. 4. Aufl. (früherer Titel: *Methods in Structural Linguistics*), Chicago 1960 (Phoenix Books. 52), 299–324.



ziehen. Konvergierende Tendenzen und der Gegensatz von Umgangs- und Literatursprache komplizieren die Verhältnisse noch weiter. Verschiedene Typen von Hilfsprachen vermitteln dort, wo scharfe Sprachgrenzen die Verständigung ganz zu vereiteln drohen. Zugleich stellt sich das Problem der Mehrsprachigkeit, worunter nach *Martinet* nur in Ausnahmefällen die vollkommene Beherrschung von zwei oder mehr Sprachen zu verstehen ist, und das der gegenseitigen Beeinflussung mehrerer Sprachen.

Das Kapitel «L'évolution des langues»<sup>13</sup> behandelt allgemein die Zusammenhänge, die bei der Untersuchung von Entwicklungen zu beachten sind. Nach *Martinet* haben sprachliche Veränderungen insofern soziale Ursachen, als sie veränderten Kommunikationsbedürfnissen entsprechen (S. 178); diese Ursachen wirken sich jedoch so langsam aus, daß inzwischen andere Ursachen auftauchen und zu wirken beginnen; bei solcher Überlagerung von Ursachen und Folgen bleibt dem Sprachwissenschaftler nichts anderes übrig, als die «inneren» Ursachen zu untersuchen, wie sie sich aus der synchronischen Beschreibung ableiten lassen, «l'étude des conflits qui existent à l'intérieur de la langue dans le cadre des besoins permanents des êtres humains qui communiquent entre eux au moyen du langage» (S. 181–182). Dazu wäre zu bemerken, daß der Beweis für eine soziale Bedingtheit sprachlicher Veränderungen, so einleuchtend sie auch scheinen mag, noch zu führen bleibt.

Ein wichtiger Ansatzpunkt der Forschung ist die Tendenz der Sprache zur Sparsamkeit, wobei zwischen *économie syntagmatique* und *économie paradigmaticque* unterschieden werden kann. Die Ökonomie findet ihre Grenze an der Notwendigkeit, eine deutliche Verständigung aufrechtzuerhalten; ob sich eine solche Grenze in dem komplexen Material einer konkreten Sprache nicht nur postulieren, sondern auch nachweisen läßt, scheint uns eines der wichtigsten Probleme der Kommunikationswissenschaft und zugleich der recht verstandenen diachronischen Sprachforschung zu sein.

Die Anwendungsmöglichkeiten der Informationstheorie auf die Sprache werden erläutert, wobei es allerdings (resignierend?) heißt: «...ce que nous devons retenir des enseignements de la théorie de l'information ressortit largement au sens commun» (S. 188). Es werden Beispiele aus Lexik, Morphologie und Phonologie gegeben. Selbst stilistische Fragen erscheinen in einem neuen Licht. Es ist Gelegenheit, daran die Gründe für die verbreitete Ablehnung quantitativer Methoden in den «Geisteswissenschaften» zu überprüfen.

Die Grenzen der quantitativen Methoden liegen allerdings dort, wo sich die Qualität der beteiligten Elemente auswirkt (S. 205). So sind Phoneme und Moneme dem angleichenden, gewissermaßen einebnenden «Einfluß» ihrer Nachbarschaft ausgesetzt (syntagmatisch), darüber hinaus aber auch einem «Systemdruck» (paradigmatisch): Es herrscht daher ein Gleichgewicht zweier Tendenzen, besser vielleicht: eine Tendenz zu ihrem Gleichgewicht. An diesem Punkt läßt sich also eine beginnende Entwicklung erfassen, wenn man über geeignete Methoden verfügt. Die möglichen Ursachen einer Weiterentwicklung schon der phonologischen Systeme können jedoch sehr mannigfaltig sein. Vermutlich empfiehlt es sich, den Begriff Ursache noch einige Zeit aus der Diskussion herauszulassen und sich bewußt mit einer vollständigen Beschreibung des Befundes und der inneren Zusammenhänge zu bescheiden. Dazu findet man reiche Anregung bei der Besprechung der Dynamik phonologischer Systeme (S. 208–217), der Übertragung funktionell wichtiger Lauteigenschaften auf benachbarte Phoneme<sup>14</sup>, der mehr oder weniger weitgehenden Integration eines Phonems in das System der *traits pertinents* und ähnlicher funktionalistischer Gesichtspunkte.

Wir haben die Besonderheiten von *Martinets* Auffassung erwähnt, ohne auf die

<sup>13</sup> Zum Folgenden vergleiche man die *Economie des changements linguistiques* (s. Anm. 7).

<sup>14</sup> Vgl. *A. Martinet: Concerning the Preservation of Useful Sound Features. Word 9: 1–11 (1953).*



Begriffsgeschichte eingehen, alle Verbindungen zu *de Saussure* und *Trubetzkoy* nachzeichnen oder abweichende Anschauungen danebenstellen zu können. Die Unterschiede zu anderen linguistischen Schulen sind zum Teil nur terminologischer Art (Morphem: Monem), zum anderen Teil methodologisch (Paradigma gegen Distribution, *traits pertinents* gegen *distinctive features*) oder schwerpunktmäßig, d. h. sie betreffen Bewertung oder Ausgestaltung von Einzelfragen (Unterscheidung von Opposition und Kontrast; marginale Stellung der Intonation). Erwähnt sei jedoch, daß *Martinets* Büchlein sich durch eine undogmatische, elastische Grundhaltung auszeichnet, die auf Schematisierung verzichtet, um neuen Einsichten gegenüber anpassungsfähig zu bleiben. In formaler Hinsicht ist die leserfreundliche Klarheit bemerkenswert. Der Text enthält dabei gerade so viel Information, daß die Aufmerksamkeit des Lesers nicht abschweift. Ein sehr nützlicher *index terminologique* (S. 221–223) sei besonders vermerkt.

Der einzige störende Druckfehler fand sich auf S. 103, wo es am Ende der letzten Zeile statt «d'unités discrètes: /ku/» heißen muß: «d'unités discrètes: /kur/».

Dem deutschen Leser seien die «*Eléments*» besonders empfohlen. Wir wünschen ihm sogar eine Übersetzung davon. Westdeutschland ist eines der letzten Länder mit sprachwissenschaftlicher Tradition, in denen der Strukturalismus fast unbekannt ist. Von vielen Linguisten und Philologen wird er ignoriert oder abgelehnt. Die Diskussionen leiden bisher nicht nur an mangelndem gutem Willen, sondern auch an unzureichender Information. Ein unerträglicher Zustand: denn viele alte Fragen bedürfen einer neuen Antwort, und viele neue Fragen sind noch nicht einmal klar formuliert. In dieser Situation kann *Martinets* realistisches und vielseitiges Buch jedem Unbefangenen eine Einführung in die moderne Sprachwissenschaft geben und jenseits aller Polemik eine neue Gesprächsgrundlage sein.

Georg Hincha, Köln

---

## Varia

---

### Ständiger Internationaler Rat – Permanent International Council Comité International Permanent

Vom 4. bis 9. September 1961 hat in Helsinki der «Vierte Internationale Kongreß für Phonetische Wissenschaften» stattgefunden. Der Kongreßbericht erscheint im Verlag Mouton & Co., Den Haag.

Während des Kongresses fanden zwei Sitzungen des Ständigen Internationalen Rats statt.

1. Sitzung am 6. September, unter Vorsitz von *E. Zwirner*, Münster. Anwesend waren *M. Cohen*, Paris; *B. Hála*, Praha; *G. E. Peterson*, Ann Arbor (Michigan); *A. Sovijärvi*, Helsinki; *F. Trojan*, Wien.

Es wurden die bisherigen Mitglieder *E. Blancquaert*, Gent; *M. P. Fouché*, Paris; *D. Jones*, Gerrard's Cross; *L. Kaiser*, Amsterdam; *A. Sommerfelt*, Oslo, zu Ehrenmitgliedern ernannt. Der Ständige Internationale Rat erweiterte sich durch Berufung von *Vl. A. Artemov*, Moskva; *P. Delattre*, Boulder; *G. Fant*, Stockholm; *D. B. Fry*, London; *R. Gsell*, Grenoble; *M. Halle*, Cambridge (Mass.); *M. Kloster Jensen*, Bergen; *B. Malmberg*, Lund; *P. Moore*, Chicago; *I. Ochiai*, Nagoya; *W. Pée*, Gent; *A. Rigault*, Montréal; *A. Rosetti*, Bucuresti; *G. Straka*, Strasbourg.

Die Herren *Håla* und *Sovijärvi* empfahlen die Wiedergründung der «Internationalen Gesellschaft für Phonetik», die im Jahre 1938 aus der Berliner «Gesellschaft für Phonetik» hervorgegangen war, sich aber infolge des Kriegsausbruches nicht offiziell konstituieren konnte.

2. Sitzung am 8. September, unter Vorsitz von *R. Jakobson*, der sein Amt niederlegte. Er und *A. de Lacerda* wurden zu Ehrenmitgliedern gewählt.

*A. Sovijärvi* berichtete über die gewählten neuen Mitglieder. Der Ständige Internationale Rat setzt sich nunmehr folgendermaßen zusammen:

*Präsident:* Professor *D. B. Fry*, London

*Vizepräsidenten:* Professor *B. Håla*, Praha; Professor *M. Halle*, Cambridge (Mass.)

*Generalsekretär:* Professor *R. Gsell*, Grenoble

*Ehrenmitglieder:* Professor *E. Blancquaert*, Gent; Professor *M. P. Fouché*, Paris; Professor *R. Jakobson*, Cambridge (Mass.); Professor *D. Jones*, Gerrard's Cross; Professor *L. Kaiser*, Amsterdam; Professor *A. de Lacerda*, Coimbra; Professor *A. Sommerfelt*, Oslo

*Mitglieder:* Professor *Vi. A. Artemov*, Moskva; Professor *S. K. Chatterji*, Calcutta; Professor *M. Cohen*, Viroflay; Professor *P. Delattre*, Boulder; Professor *G. Fant*, Stockholm; Professor *E. Fischer-Jørgensen*, København; Dozent *M. Kloster Jensen*, Bergen; Professor *B. Malmberg*, Lund; Professor *P. Moore*, Chicago; Professor *I. Ochiai*, Nagoya; Professor *W. Pée*, Gent; Professor *G. E. Peterson*, Ann Arbor (Mich.); Professor *A. Rigault*, Montréal; Professor *A. Rosetti*, Bucuresti; Professor *A. Sovijärvi*, Helsinki; Professor *G. Straka*, Strasbourg; Professor *F. Trojan*, Wien; Professor *E. Zwirner*, Münster.

Professor *Zwirner* wurde zum Präsidenten des 5. Internationalen Kongresses 1964 gewählt.

Die Wahlen beider Sitzungen erfolgten einstimmig.

Es wurde vorgeschlagen, *B. Malmberg*, Lund, als Präsidenten der neu zu gründenden «Internationalen Gesellschaft für Phonetik» zu empfehlen und ihn zu bitten, sich seinen Generalsekretär selbst zu bestellen und die Gründung der Gesellschaft vorzubereiten.

*R. Jakobson* riet zu einer Autonomie der «Gesellschaft» und des «Ständigen Internationalen Rats». Dem Vorschlag wurde zugestimmt.

Es wurde beschlossen, das Protokoll der Sitzung in den «Phonetica» zu veröffentlichen.

*E. Z.*

## Fourth International Congress of Phonetic Sciences

The increasing importance of research in the phonetic sciences was effectively emphasized at the most recent Congress. The previous congresses were held in Amsterdam, London, and Ghent, in 1932, 1935, and 1938, respectively. The following is a brief report of the Fourth International Congress of Phonetic Sciences which was held in Helsinki, September 4–9, 1961. There were 240 scholars, representing 23 different countries. Most of the leading phoneticians from America and Europe attended the Congress.

The Congress was organized into plenary session and section meetings. At the plenary sessions, invited papers were presented, primarily to survey research progress in phonetics since the previous congresses. Most of the papers presented at the section meetings were concerned with reports of specific research investigations. During the first day of the Congress the plenary session extended throughout most of the day, and on successive days it extended until noon, with section meetings following throughout the afternoon.



The plenary sessions on the first two days of the Congress were devoted to acoustic and physiological phonetics. On Wednesday the psychological aspects of phonetics were discussed, and the last two plenary sessions were devoted to phonetics and phonemics. Altogether, a total of 25 papers were planned for the plenary sessions, and with a few exceptions the papers were presented by the authors as listed in the program.

A review of the section meetings shows a heavy emphasis upon physiological and acoustic phonetics. There were also a number of papers concerning the relation of phonetics to the psychological and linguistic aspects of language. Several scholars who had submitted papers for the section meetings were unable to attend, and as a result sessions were combined in some cases. As a whole, 72 papers were read. The many new ideas that were presented, and the general good attendance at the Congress contributed to much active discussion at the section meetings. The fine program of scientific and pedagogical films also should be noted.

Phonetics was identified at the Congress as a scientific discipline basic to many areas, including speech pathology, foreign language teaching, experimental and theoretical linguistics, and speech analysis and synthesis for technical communication purposes. Papers applying phonetics to all of these areas were presented, and those with a technical orientation appeared to be generally best received. Modern techniques and research areas were clearly in evidence, such as high speed laryngeal photography, X-ray laminagraphy and cinematography, subglottal pressure measurements, sound spectrography, electrical vocal analogs, speech synthesis, and automatic speech recognition.

In this brief report it would be impossible to enumerate the many specific contributions to phonetics and phonemics which were presented at the Congress. However, special note must be taken of the excellent presentations by the President of the Congress, the President of the Permanent Council, and by the new presidents who have now taken office. The four papers are as follows:

Professor *Antti Sovijärvi*, President of the Fourth International Congress

Röntgenkinematographisch-akustische Untersuchungen über die Artikulation der Diphthonge

Professor *Roman Jakobson*, formerly President of the Permanent Council  
The Phonemic Concept of Distinctive Features

Professor *Eberhard Zwirner*, President of the Fifth International Congress  
Beitrag zur Geographie der prosodischen Eigenschaften

Professor *Dennis B. Fry*, now President of the Permanent Council  
Automatic Speech Recognition

For those who came to the Congress from abroad, there were also rich cultural and social opportunities. There were receptions, visits to scientific institutes throughout Helsinki, and an afternoon excursion by boat to Porvoo. The day after the closing sessions of the Congress the members traveled to Turku, where the Cathedral, Abo Akademi, and Turku University were visited. A banquet had been arranged for the members of the Congress within the beautifully restored great hall of the Castle at Turku. It was at this banquet that several delegates to the Congress took the opportunity to express the appreciation of all of those who attended to Professor *Antti Sovijärvi*, President, and Professor *Pentti Aalto*, Secretary of the Congress, for the excellent manner in which the entire meeting had been arranged and conducted. These men are indeed to be congratulated for having provided a unique opportunity for those interested in phonetics to meet together in the common purpose of developing an increased understanding of the scientific aspects of human communication.

The *Proceedings* of the Fourth International Congress are now being prepared for publication. The Fifth International Congress is to be held in Bochum, Germany, in 1964, with Professor *Eberhard Zwirner* as President of the Congress.

*Gordon E. Peterson*, Ann Arbor, Mich.

# Zeitschrift für deutsche Wortforschung

Begründet von  
FRIEDRICH KLUGE

In Verbindung mit  
RALPH FARRELL, EMIL ÖHMANN, ERICH ROTHACKER, OTTO SPRINGER, WILHELM WISSMANN

Herausgegeben von  
WERNER BETZ

Mit dem sechzehnten Band, dem ersten der Neuen Folge, nimmt die ZEITSCHRIFT FÜR DEUTSCHE WORTFORSCHUNG ihre 1914 unterbrochene Tradition wieder auf. Es ist vorgesehen, daß jährlich drei Hefte im Umfang von je vier Bogen erscheinen, es können auch zwei Hefte zu einem Doppelheft zusammengefaßt werden. Das Format ist gegenüber den alten Bänden vergrößert worden.

Der Preis des Jahrganges beträgt DM 24.—.

Einzelhefte werden nicht abgegeben. Die Zeitschrift kann nur zur Fortsetzung geliefert werden. Einbanddecken für je DM 2.50 werden jeweils zusammen mit dem Schlußheft ausgegeben.

## **Inhalt von Band 16, Heft 1/3 (Band 1 der neuen Folge)**

WERNER BETZ, Zur neuen Folge

### **I. Aufsätze**

ERICH ROTHACKER, Das Wort «Historismus»; FRITZ TSCHIRCH, Bedeutungswandel im Deutsch des 19. Jahrhunderts; ERIK ERÄMETSÄ, Adam Smith als Mittler englisch-deutscher Spracheinflüsse; HARRI MEIER, Das lateinische V und die alten Germanen; LOUIS L. HAMMERICH, Über die Modalverba der neugermanischen Sprachen (mit besonderer Berücksichtigung des Dänischen); CORNELIA BERNING, Die Sprache des Nationalsozialismus; MARGA MEHRING, Menschenmaterial; PAUL ZINSLI, «Im Tuppuwald»; EMIL ÖHMANN, Die deutsche Wortsippe «Hast»; HERBERT KOLB, Der inhumane Akkusativ;

### **II. Kleine Beiträge**

WERNER BETZ, Zum «Wörterbuch des Unmenschen»; HANS SCHMEJA, Zum Wort «Phase»

## **Inhalt von Band 17, Heft 1/2 (Band 2 der Neuen Folge)**

### **I. Aufsätze**

WOLFGANG SCHLACHTER, Ein Aktionskriterium im Neuhochdeutschen; KEITH SPALDING, Bemerkungen zum Eisernen Vorhang; FRIEDRICH NEUMANN, Einstellung als seelisches Verhalten; EMIL PLOSS, Stanniol; CORNELIA BERNING, Die Sprache des Nationalsozialismus (Fortsetzung)

### **II. Kleine Beiträge**

ROLF GLUNK, Knüller; GERHARD EIS, *Verpitschieren* in der altdeutschen Weinbau- und Rennterminologie



**Walter de Gruyter & Co.**

**Berlin W 30**

vormals G. J. Göschen'sche Verlagshandlung – J. Guttentag, Verlagsbuchhandlung – Georg Reimer – Karl J. Trübner – Veit & Comp.



# Die «Sprache» der Tiere

**Tatsachen – Problemschau – Theorie**

**Von Professor Dr. Fr. Kainz**

Universität Wien

*1961. VIII, 322 Seiten. Geheftet DM 26.–, Ganzleinen DM 29.50*

Der bekannte Autor des fünfbandigen Werkes «Psychologie der Sprache» bietet in diesem Buch eine auf breitem zoologischem (ethologisch-ökologischem) Tatsachenfundament errichtete Erörterung spezieller Probleme der Tierpsychologie und vergleichenden Verhaltensforschung, die in unseren Tagen wachsende Bedeutung gewonnen und damit zunehmendes Interesse gefunden haben. Auf der Basis der von den genannten Disziplinen gewonnenen Erkenntnisse wird eine vielseitige Theorie der animalischen Kontaktprozeduren entworfen und durch Folgerungen aus eigenen Beobachtungen und Experimenten ergänzt. Ein besonderes Anliegen des Verfassers ist hierbei die Darstellung systematisch-prinzipieller Konsequenzen, die sich durch diese Forschungen vom Andersartigen und dem Kontrast her für die Sprache des Menschen ergeben.

Der erste Teil des Werkes enthält in methodischer Gliederung das vielfältige biologische Tatsachenmaterial; der zweite Abschnitt bringt auf dem Weg wechselseitiger Erhellung sowohl die Eigenart der Tierkommunikation als auch die Wesensbesonderheit der Sprache überhaupt in neuer Sicht. Das Buch greift somit in den noch nicht entschiedenen «Kampf um die Tier-sprache» ein und bringt einen wesentlichen Beitrag zur Entscheidung dieser Auseinandersetzung. Das Verfahren, mit dem dieser Versuch unternommen wird, ist das einer behutsam vorgehenden Problemdiskussion, die jedes Für und Wider sorgsam abwägt und die zentralen Fragen von verschiedenen Seiten aus und unter mannigfachen Einbettungsaspekten darlegt.

Schließlich findet auch die tiersoziologische Problematik gebührende Berücksichtigung, da im Zusammenhang tiersoziologischer Erörterungen vielfach Feststellungen getroffen werden, denen auch sprachtheoretische Bedeutung zukommt.

---

FERDINAND ENKE VERLAG      STUTTGART